2407

Value states and decision making in the prisoner's dilemma game

A. L. M. van der Sanden

ERRATA

- *The title on the cover should read:* Value states and dynamic decision making in the prisoner's dilemma game
- *page 4, line 2:* bij → by
- *page 19, line 30:* simular on → similar to
- *page 23, line 2:* strrategy → strategy
- *page 23, line 4:* substituding → substituting
- *page 23, line 22:* ar (R_1,C_1) → at (R_1,C_2)
- *page 23, line 29:* nonzer-sum → nonzero-sum
- *page 24, line 1:* $C_2 \rightarrow C_1$
- *page 26, lines 8 to 12: The phrase starting with* "It appears that" *must be inserted after line 4 of page 25.*
- *page 26, line 21:* $(R_1,C_2) \rightarrow (R_1,C_1)$ *and* $(R_2,C_1) \rightarrow (R_2,C_2)$
- *page 33, line 13 after Table 1.1:* expections → exceptions
- *page 38, line 9 after Table 1.2:* risk decision → risky decision
- *page 38, line 12 after Table 1.2:* greet → greed
- *page 41, line 11:* in *C* and *D* → in two value states, *C* and *D*,
- *page 54, line 2 after Table 2.1:* preferrt → preferred
- *page 54, line 3 after Table 2.1:* payer → player
- *page 55, line 24:* 1.e. → i.c.
- *page 61, line 18 after Table 2.5:* $z = a \rightarrow z = -a$
- *page 69, line 20:* (2.28) → (2.18)
- *page 72, line 13:* way which → way in which
- *page 77, line 6 after Table 2.14:* sence → sense
- *page 80, line 6:* Wit → With
- *page 88, line 16:* *Axiom* 3-2 → *Axiom* 3-3
- *page 93, line 2:* in → is
- *page 96, line 30:* in → is
- *page 105, line 4:* assingment → assignment
- *page 115, line 13:* $r_j \rightarrow r_i$
- *page 119, last line of Table 4.2:* *d* or *c*, resp., → *c* or *d*, resp..
- *page 122, line 7:* prodect-set → product-set
- *page 124, line 2:* $(d,d/c) \rightarrow (d,d/d)$
- *page 124, line 3:* $\{(cd), (dd), (cc)\} \rightarrow \{(cd), (dd)\}$
- *page 124, line 15: insert* "$\Rightarrow$" *between* "$(h(f),f)$" *and* "$(h(f), f(h(f)))$"
- *page 125, line 30:* P_1 and P_2, such that $UP_1 \rightarrow p_1$ and p_2, such that Up_1

ERRATA (continued)

- *page 130, line 25:* $(p_1,f_1) \rightarrow (p_i,f_i)$
- *page 130, line 26:* $p_1 \rightarrow p_i$ *and* $f_1 \rightarrow f_i$
- *page 132, line 6:* If $\bar{}$ is $\rightarrow$ If $\bar{a}$ is
- *page 133, lines 1 and 5:* i-commitment $\rightarrow$ i-metacommitment
- *page 138, line 20:* wath $\rightarrow$ what
- *page 142, line 1:* in $\rightarrow$ is
- *page 142, line 32:* *tranform* $\rightarrow$ *transform*
- *page 151, line 9:* were $\rightarrow$ was
- *page 162, lines 18 and 22:* Shell $\rightarrow$ Snell
- *page 163, line 18:* antries $\rightarrow$ entries
- *page 176, line 2 after Table 5.8:* (see equation 5.5) $\rightarrow$ (see equation 5.6)
- *page 178, line 8:* indentified $\rightarrow$ identified
- *page 178, line 31: insert* "and" *between* "state" *and* "assuming"
- *page 179, line 12 should read:* all j on which event X occurs, m_2 the *etcetera*
- *page 180, line 16 should read:* all k on which event Y occurs, m_2 the *etcetera*
- *page 181, line 3:* likelohood fuction $\rightarrow$ likelihood function
- *page 181, line 6:* (2m +m +m) $\rightarrow$ $(2m_1+m_2+m_3)$
- *page 181, line 12:* v_2 to $v_2 \rightarrow v_1$ to v_2
- *page 184, line 10: delete* "on play j"
- *page 184, lines 17/18: delete* "on play k"
- *page 185, line 18:* centext $\rightarrow$ context
- *page 186, line 16:* erea $\rightarrow$ area
- *page 187, line 2:* Indentifiable $\rightarrow$ Identifiable
- *page 187, line 3:* All-or-one $\rightarrow$ All-or-none
- *page 188, line 25:* *matagames* $\rightarrow$ *metagames*

-o-o-o-

concerning: corrections dissertation
"Value states and dynamic decision
making in the prisoner's dilemma game"
by A.L.M. van der Sanden

Brunssum, October 25, 1979

L.S.,

please, insert the enclosed corrections in my dissertation you received some days ago.

Yours sincerely,

A.L.M. van der Sanden
Vondelstraat 36
6445 AK Brunssum, The Netherlands

VALUE STATES AND DYNAMIC DECISION MAKING IN THE PRISONER'S DILEMMA GAME

"Nessuna certezza è dove non si pò applicare una delle scienzie matematiche, ovver che non sono unite con esse matematiche."

(Leonardo da Vinci, 1452-1519)

Promotor: Prof. Dr. E.E.Ch. I. Roskam

VALUE STATES AND DYNAMIC DECISION MAKING IN THE PRISONER'S DILEMMA GAME

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR
IN DE SOCIALE WETENSCHAPPEN
AAN DE KATHOLIEKE UNIVERSITEIT TE NIJMEGEN,
OP GEZAG VAN DE RECTOR MAGNIFICUS
PROF. DR. P.G.A.B. WIJDEVELD,
VOLGENS HET BESLUIT VAN HET COLLEGE VAN DECANEN
IN HET OPENBAAR TE VERDEDIGEN OP
DONDERDAG 15 NOVEMBER 1979
DES NAMIDDAGS TE 4.00 UUR

DOOR

ANTONIUS LAMBERTUS MARIA VAN DER SANDEN

GEBOREN TE 'S-HERTOGENBOSCH

1979
DRUK: STICHTING STUDENTENPERS NIJMEGEN

PREFACE

This study was supported bij a grant from The Netherlands Organization for the Advancement of Pure Research (ZWO), project number 56-60. The experiments reported were carried out while I was at the Department of Psychology of the Catholic University of Nijmegen.

Parts of this dissertation were published before. Chapter 1 is entirely new. The second chapter was published before in *The British Journal of Mathematical and Statistical Psychology*. Chapter 3 is a thoroughly revised and extended adaptation of Technical Report 76 MA 02 of the Department of Psychology, University of Nijmegen. Chapter 4 is a reprint of Technical Report 79 MA 01 of the Department of Psychology, University of Nijmegen. Chapter 5 is adapted from an earlier draft written as Technical Report 79 MA 02 of the Department of Psychology, University of Nijmegen.

Many people, colleagues and students, have contributed to this study either materially or through their comments and criticisms or through their being a subject in one of the experiments reported. I would like to express my gratitude to all of them.

I want to apologize to my colleagues of the Central Bureau of Statistics for the neglect they suffered because of my Ph.D.-work.

Finally, I want to express my indebtedness to my wife. Without her continuous encouragements, probably, this dissertation would not have been completed yet.

Aan Annelies,
aan Pieter en Judith

CONTENTS

1.1. *Introduction*

When two parties with partly coincident and partly opposed interests are interacting, one speaks of a social conflict. Each party may consist of a single actor or of a group of actors with common interests. More generally, the problem of cooperation and competition in these situations of conflict has been studied quite extensively by many researchers in the behavioral sciences.

Since the epoch-making book of Von Neumann and Morgenstern, *Theory of games and economic behavior*, was published in 1944, human game playing behavior has been a significant area of research in the social sciences.

The theory of games is a mathematical system of definitions and theorems about decision situations, called games, in which there are $n \geqslant 2$ "players" who choose from a set of strategies. A strategy is essentially a plan chosen by a player in which the player's choice among available alternatives is indicated in every situation that can arise during a play of the game (Rapoport, 1970). The game is played by having each player choose one strategy. The joint strategy choice of all players involved determines the outcome of the game.

The conflict which is embedded in a game, arises from the fact, that different players in the game have different preference orders for the outcomes.

If there are in fact only two players ($n = 2$), and if each player has a finite number of strategies, then this situation may be represented as in Figure 1.1 by a matrix whose rows are one player's strategies, denoted $R_1, \ldots, R_k$, and whose columns are the other player's strategies, denoted $C_1, \ldots, C_m$. The cell entries represent the outcomes of the game. More specifically, O_{ij} denotes the outcome that results of row player choosing R_i and column player choosing C_j.

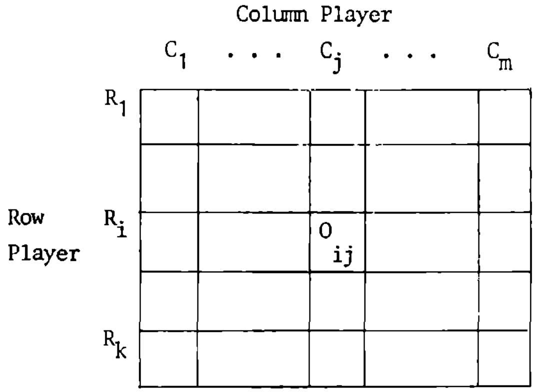


Figure 1.1 Two-person game in normal form.

In general, payoffs are attached to each outcome, one for each player. The payoffs attached to O_{ij} will be indicated as r_{ij} and c_{ij} for row player and colomn player respectively. This actually defines what is called a game in *normal form*.

The structure that underlies the game in normal form is the game in extensive form. A game in extensive form (or a game tree) is a representation of the scenario of the game in normal form as it is actually played. To illustrate this we consider the game called Button-Button (see Rapoport, 1966a, p.65-68).

Button-Button is played as follows. *Hider* conceals a button in one of his hands, and *Guesser* tries to guess in which hand the button is concealed. The game has four outcomes: (1) button in left hand and *Guesser* says "Left"; (2) button in left hand and *Guesser* says "Right"; (3) button in right hand and *Guesser* says "Left"; (4) button in right hand and *Guesser* says "Right".

Suppose *Hider's* choice is known to *Guesser* and the payoffs are different for both players depending on the outcome. Now the game can be represented by a tree diagram that describes the available moves to each player at each stage of the game. This game tree is shown in Figure 1.2.

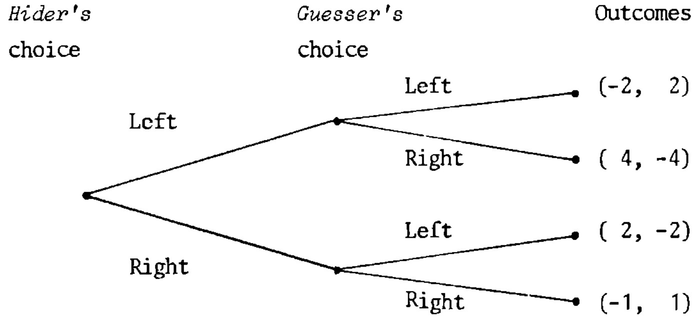


Figure 1.2 Button-Button in extensive form. The payoff to *Hider* is mentioned first, and the payoff to *Guesser* is mentioned second in the braces.

To reduce this to a game in normal form we must assign four strategies to *Guesser* (to do justice to the asymmetry that exist because *Hider's* choice comes first). A strategy is now a complete plan of action covering all contingencies.

The strategies of *Hider* are: H_1 (left) and H_2 (right). *Guesser's* strategies are: G_1 (guess left regardless of where the button is), G_2 (guess the hand where the button is), G_3 (guess the hand where the button is not), and G_4 (guess right regardless of where the button is): Now we obtain the normal form shown in Fig. 1.3.

		Guesser			
		G_1	G_2	G_3	G_4
Hider	H_2	-2	-2	4	4
	H_2	2	-1	2	-1

Figure 1.3 Normal form of the extensive form of Figure 1.2. The payoff to *Guesser* equals the payoff to *Hider* with reversed sign (*Hider's* payoffs are displayed).

In the absence of the information as to where the button is, *Guesser* can only guess unconditionally (left or right), and the game becomes as is shown in Figure 1.4.

		Guesser	
		G'_1	G'_2
Hider	H_1	-2	4
	H_2	2	-1

Figure 1.4 Button-Button in normal form.
G'_1 and G'_2 are unconditional strategies. Each player chooses without knowledge of the other player's choice. (G'_1 = "left"; G'_2 = "right").

Experimental games may have only a single play or may be played a number of times (trails) in succesion by the same set of players. In the latter case one speaks of *iterated plays* of the game. Experiments on iterated plays of a game are very common in behavioral research. They offer the oppertunity to study the sequential characteristics of game playing behavior, that is players will play with memory of past iterations and will anticipate future ones.

1.2. *Game theory as a normative theory of rational behavior*

Originally developed as a special branch of the theory of decision making game theory is defined as a theory of rational behavior. This implies that it is assumed, that each player in a game is individually rational, in the sense that his preference ordening of the outcomes is determined by the order of magnitude of the associated payoffs to him and that he is rational in the sense that he assumes that every other player in the game is also individually rational (Rapoport, 1974).

The aim of game theory is to prescribe decision policies appropriate to conflict situations, formalized as games, in such a way that the policies are seen as necessary, logical consequences derived from the assumptions of rationality and from the constraints of the situation (Rapoport and Orwant, 1962). Thus game theory is normative rather than descriptive, that is, its conclusions state how "rational" people ought to behave rather than how real people do behave.

When applied to different types of games the concept of rationality becomes rather vaque. This is demonstrated in this section. Some important categories of two-person games and their "solutions" are discussed. Examples of games are given for wich different solution concepts all based on the principle of rationality yield contradictory results.

Unambiguous priscriptions based on the principle of rationality are easily derived as long as two-person zero-sum games are treated. A two-person game is said to be zero-sum (or strictly competitive) if, for any outcome of the game, the payoff to one player is exactly balanced by the payoff to the other player, that is, the algebraic sum of the payoffs to each player is always zero. If this sum is always constant, not necessarily zero, regardless of the outcome of the game, the game is called "two-person constant-sum". Constant-sum games are not essentially different from zero-sum games.

A *dominating strategy* is a strategy which is preferred by some player over all other available strategies. If a two-person zero-sum game has a dominating strategy for any player, it is prescribed by the criterion of rationality. Otherwise, the minimax criterion leads to a choice of strategy which guarantees the greatest possible win under the constraints of the game (Rapoport and Orwant, 1962). The *minimax* criterion prescribes a player to choose the strategy for which the minimum attainable payoff is maximum.

1.2.1. *Two-person zero-sum games with saddle points*

Zero-sum games are games where the interests of the players are diametrically opposed. In such games one player's gain is his opponent's loss.[1)]

A saddle point in the matrix of a two-person zero-sum game is an outcome where the payoff is at the same time the smallest in its row and the largest in its column. (If the game is zero-sum, only row player's payoffs are entered since the other payoff must be the same number with the opposite sign.)

		Column player		
		C_1	C_2	C_3
	R_1	-7	-5	7
Row player	R_2	10	-2	-5
	R_3	3	1	4

Figure 1.5 Two-person zero-sum game with saddle points; entries are Row player's pay-offs (gains).

In Figure 1.5 the saddle point is in row R_3 and column C_2. It is the smallest entry in its row and the largest in its column.

The choice of R_3 by the row player is called the *maximin strategy*, because it maximizes the minimal value that can be obtained in a row. The choice of C_2 by the column player is called the *minimax strategy*, because it minimizes the maximum value that he may lose in a column.

The strategies R_3 and C_2 are in *equilibrium* because it is not advantageous to either player to change his strategy as long as his opponent does not change his strategy.

1) The assumption of strict opposition does not necessarily apply to the numerical utilities of the payoffs. However, this problem will be ignored here.

There may be more than one point in a single zero-sum game, that are minimum in its row and maximum in its colomn. It can be shown that if a game-matrix has several saddle points, the corresponding payoffs are all equal; moreover, the coordinates of these saddle points are interchangeable. That is, if O_{ij} and O_{kl} are both saddle points, then O_{il} and O_{kj} are also saddle points with equal entries.

The concept of "best" or "rational" strategy is so straightforward and convincing in zero-sum games with saddle points, that maximin and minimax strategies are the only plausible choices. Therefore, the normative aspect of game theory is strongest with respect to zero-sum games. At the same time this delimits the use of game theory. As Rapoport and Orwant put it (1962, p.4):"...*an experiment on a game with saddle points will (...) be guided by a hypothesis that the saddle-point strategies will in fact be chosen by sufficiently intelligent players. If the results confirm the hypothesis, little remains to be asked further.*"

1.2.2 *Two-person zero-sum games without saddle points*

It is easy to show that in zero-sum games a saddle point exists if and only if the maximin and the minimax strategies are in equilibrium. Many games do not have a saddle point, and hence no pure stategies that are in equilibrium.

If the game matrix has no saddle point, there exists for each player an optimal *mixed strategy*, that is, a probability distribution on the available strategies, which has the property that the resulting *expected values* of the gains to each player respectively have minimax properties.[1)]

In the game matrix of Figure 1.6 there is no saddle point, because no entry which is minimum in its row is at the same time maximum in its column. The maximin and minimax strategies are R_1 and C_1 respectively.

1) Strictly speaking one should say "maximin properties", because the mixed strategy maximizes the minimum expectation.

		Column player	
		C_1	C_2
Row player	R_1	-1	6
	R_2	3	-4

Figure 1.6 Two-person zero-sum game without saddle points

However, if Row player expects Column player to choose C_1, Row player will select R_2 because that will yield a more preferred payoff. If this is anticipated by Column player, Column player is better off when selecting C_2. But now, if Row player expects Column player to select C_2, Row player should choose R_1. And so on ad infinitum...

The game theoretical prescription in this case is for Row player to select R_1 and R_2 with equal probabilities (.5), and for Column player to select C_1 with probablility 5/7 and C_2 with probability 2/7.

If p_1 and q_1 are the probabilities of selecting R_1 and C_1 respectively, the optimal mixed strategies can be easily found (see, e.g. Coombs, Dawes and Tversky (1970)) as

$$p_1 = \frac{r_{22} - r_{21}}{r_{11} + r_{22} - r_{12} - r_{21}}$$

$$q_1 = \frac{r_{22} - r_{12}}{r_{11} + r_{22} - r_{12} - r_{21}}$$

The expected value of the gains that each player can secure, regardless of the strategy of his opponent, indicated as v, is obtained as

$$v = \frac{r_{11} * r_{22} - r_{12} * r_{21}}{r_{11} + r_{22} - r_{12} - r_{21}}$$

The value of the game in Figure 1.6 equals 1.

The rational behind the optimal mixed strategy (or minimax mixed strategy) is appealing, when the game is played repeatedly and payoffs can be cummulated. Then the expected payoff will be realized as the long-run average payoff.

It may be clear, that in two-person zero-sum games without saddle points, departures from solutions prescribed by the normative theory will be found more frequently than in two-person zero-sum games with saddle points.

First, even if the principle of minimax mixed strategies is known to the player of the game, the calculation of such a strategy is a complex computational task. Players ignorant of game theory cannot be expected to accomplish such a task.

Second, knowledge about the utility functions on the payoffs is required to compute or to evaluate mixed strategies. In the case of games with a saddle point only a preference order among the outcomes needs to be determined, because the saddle point is invariant with respect to any monotone tranformation of the payoffs. Tranformations more general than the linear transformation may affect the result of the averaging procedure involved in calculating a mixed strategy.

1.2.3. *Two-person nonzero-sum games*

Minimax is a clear-cut criterion for zero-sum games. That is, it can be defended on logical and rational grounds and accepted as "optimal behavior"[1] for this class of games. Problems arise when nonzero-sum games are considered. Here the rationality criterion breaks down.

In a nonzero-sum game for at least one outcome the algebraic sum of the payoffs to each player is not zero. A different name for nonzero-sum games is *partly competitive games*, because the interests of the players are not diametrically opposed. In these games ambivalence about the "best" decision is introduced, because different criteria based on the principle of rationality lead to

1) "optimal behavior" in the sense of maximizing expected payoff.

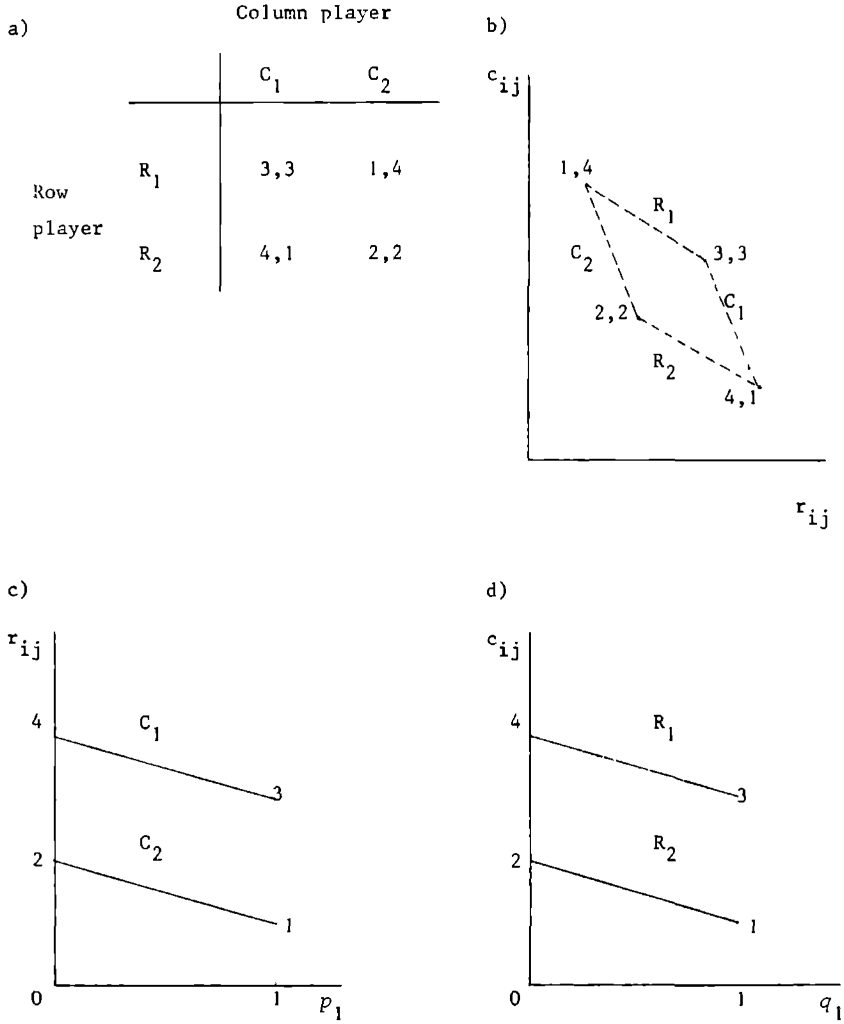


Figure 1.7 Prisoner's Dilemma Game. In part (a) the game matrix is given. Part (b) is a graphical representation of the payoff structure; Row's payoffs are indicated on the absciss and Column's payoffs are indicated on the ordinate. In part (c) Row's expected payoffs are displayed as a function of his probability of selecting R_1 for both pure strategies of Column. Part (d) is equivalent to (c) except that the roles of Row and Column are reversed.

different prescriptions.

Nonzero-sum games are called *cooperative* or *negotiable* when the outcome can be negotiated by the players (Coombs *et al.*, 1970). Otherwise, the games are called *noncooperative* or *nonnegotiable*. This is a distinction between the conditions under which the games are played. In this section we shall discuss some nonzero-sum games, which are all noncooperative.

A prominent example is the *Prisoner's Dilemma Game* (PDG), which is a two-person nonzero-sum game where, if both players choose their dominating (i.e. rational) strategy, each gets less than if either one had chosen his dominated strategy (see Figure 1.7).

The principal feature of PDG is that for both players the second strategy is a dominating strategy, that is, R_2 dominates R_1 for Row player and C_2 dominates C_1 for Column player. However, the choice of (R_2, C_2) results in a non-optimal outcome. The outcome (R_1, C_1) is preferred to (R_2, C_2) by both players. This is also illustrated in the graphical representation of the game in Figure 1.7 (b, c, d). The outcome (R_2, C_2) is an equilibrium (because no player can improve his payoff by unilaterally changing his strategy) and satisfies the minimax principle.

Another example of a two-person nonzero-sum game is the game called "*The inspector and the thief*" (Figure 1.8). In "*The inspector and the thief*" R_1 and C_1 are pure minimax strategies, but although R_1 is also optimal for Row player, the optimal strategy for Column player is a mixed strategy, $q = (q_1, 1-c_1)$, with $q_1 = 3/4$. This value is found as:

$$q_1 * c_{11} + (1-q_1) * c_{12} = q_1 * c_{21} + (1-q_1) * c_{22},$$

which yields in the present case

$$3q_1 + (1-q_1) = 2q_1 + 4(1-q_1) \Longrightarrow q_1 = 3/4$$

This derivation is simular on the one given for the computation of optimal mixed strategies for zero-sum games without saddle points (cfr. page 16).

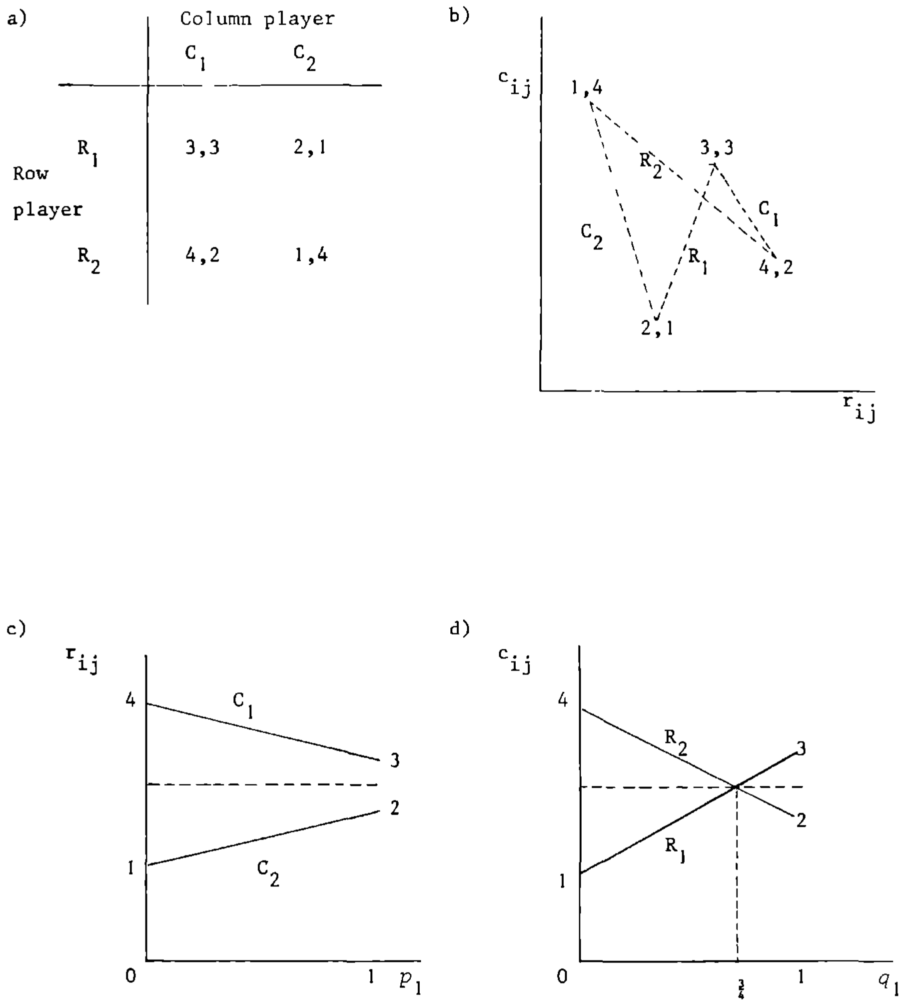


Figure 1.8 The inspector and the thief

(R_1,C_1) is not an equilibrium, because if Row player expects C_1, he will select R_2. But if this is foreseen by Column player, he will choose C_2, etcetera.

Similarly, it is easy to see that the optimal strategies are not in equilibrium. If Row player chooses R_1 with probability 1, the best thing for Column player to do is to select C_1 with probability 1. This brings us back to the circular reasoning of the last paragraph.

If a game does not have a dominating strategy for each player, mixed strategies can be computed, which yield an equilibrium. That is these mixed equilibrium strategies yield expected gains such that any player cannot improve his expected gains by playing a different (mixed) strategy as long as the other players stick to these mixed equilibrium strategies.

Let v_r be the expected gain for Row player regardless of his own (mixed) strategy and let v_c be the expected gain for Column player regardless of his own (mixed) strategy. (Note the difference with the definition of v in the case of zero-sum games without a saddle point on page 16!). Now we derive :

$$v_r = p_1\ (q_1*r_{11} + (1-q_1)*r_{12}) + (1-p_1)*(q_1*r_{21} + (1-q_1)*r_{22})$$

$$= p_1\ ((r_{12}-r_{22}) - q_1\ (r_{12}-r_{22}+r_{21}-r_{11})) + (r_{22}+q_1*(r_{21}-r_{22})),$$

which is independent of p_1, for q_1, equal to

$$q_1^* = \frac{r_{12} - r_{22}}{r_{12} - r_{22} + r_{21} - r_{11}} \quad ,$$

and $v_c = q_1\ (p_1*c_{11} + (1-p_1)*c_{21}) + (1-q_1)*(p_1*c_{12} + (1-p_1)*c_{22})$

$$= q_1\ ((c_{21}-c_{22}) - p_1(c_{21}-c_{22}+c_{12}-c_{11})) + (c_{22}+p_1*(c_{12}-c_{22})),$$

which is independent of q_1 for p_1 equal to

$$p_1^* = \frac{c_{21} - c_{22}}{c_{21} - c_{22} + c_{12} - c_{11}}$$

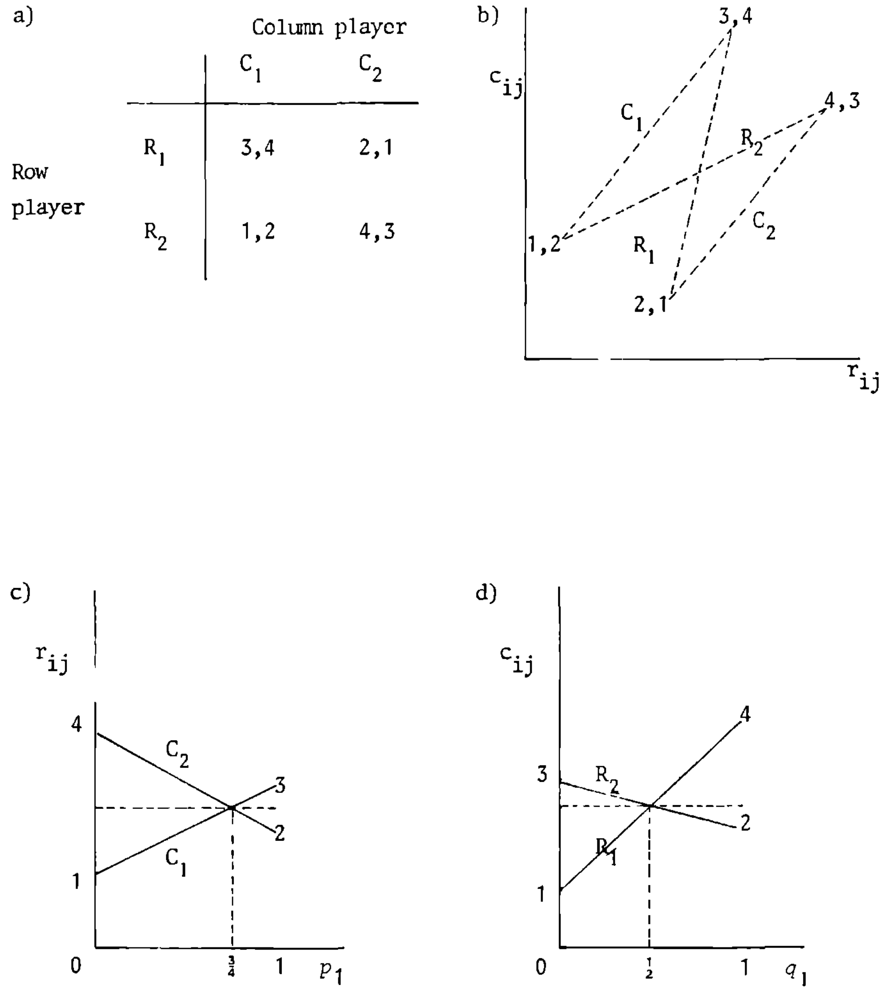


Figure 1.9 Lucas and Mattheus

These strategies, p_1^* and q_1^*, are equilibrium strategies, because if any player's opponent follows the equilibrium strrategy, this player cannot improve his expected payoff. The values of v_r and v_c can be readily found by substituding the expressions of p_1^* and q_1^* for p_1 and q_1 respectively.

This yields

$$v_r = \frac{r_{12} * r_{21} - r_{11} * r_{22}}{r_{12} - r_{22} + r_{21} - r_{11}}$$

and

$$v_c = \frac{c_{12} * c_{21} - c_{11} * c_{22}}{c_{21} - c_{22} + c_{12} - c_{11}}$$

In this case of "*The inspector and the thief*" the equilibrium strategies are: $p_1^* = \frac{1}{2}$ and $q_1^* = \frac{1}{2}$. The corresponding equilibrium values are: $v_r = 2.5$ and $v_c = 2.5$.

In this game the pure minimax strategy, (R_1, C_1), yields a payoff which is preferred to the (expected) payoff of the equilibrium strategy, but the minimax strategy is not equilibrium.

Consider the game in Figure 1.9. This game is called "*Lucas and Mattheus*". The pure minimax strategies are R_1 and C_1, which yield an equilibrium. There is also a second equilibrium, namely (R_2, C_2).

Each equilibrium is optimal for one player and nonoptimal for the other one. If both players select their strategy so as to obtain the equilibrium with the optimal payoff for the opponent (altruistic choice), they arrive ar (R_1, C_1), which is worst or next to worst. Similarly, if both players select their strategy so as to obtain the equilibrium with the optimal payoff for oneself (individualistic choice), they arrive at (R_2, C_1), which is also worst or next to worst.

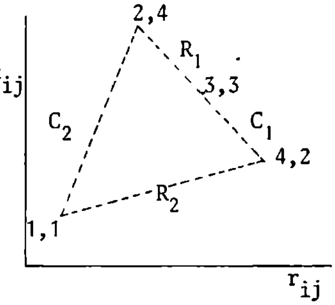
Again the pure minimax strategy, (R_1, C_1), yields a payoff which is preferred to the expected payoff of the mixed equilibrium strategy.

A nonzer-sum game which very much looks like the PDG is "*Game of Chicken*" (Figure 1.10). There is not a dominating strategy for any player in this game. R_1 and C_2 are pure minimax (and optimal mixed) strategies, although not in equilibrium. (R_2, C_1) and (R_1, C_2) are both equilibrium. If either player can bluff the other that he will select his second strategy, the other will do best to select his first strategy, and the player who bluffs will obtain his best payoff.

a)

Row player	Column player: C_1	Column player: C_2
R_1	3,3	2,4
R_2	4,2	1,1

b)



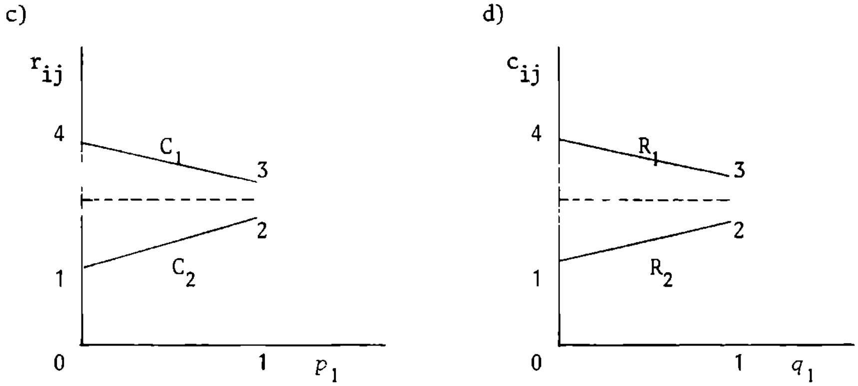


Figure 1.10 The Game of Chicken

But if both bluff each other, and carry out their bluffs, each player gets his worst payoff.

The mixed equilibrium strategies for the "*Game of Chicken*" turn out to be: $p_1^* = \frac{1}{2}$ and $q_1^* = \frac{1}{2}$, with expected gains $v_r = v_c = 2.5$.

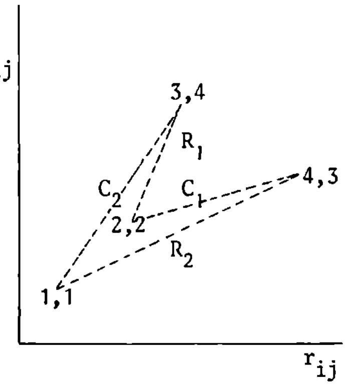
A slight modification of the game matrix of Figure 1.10 yields the game known as "*The Battle of the Sexes*" (or sometimes called "*Mild Chicken*"). This game is shown in Figure 1.11.

In "*The Battle of Sexes*", like in "*The Game of Chicken*", there is no dominating strategy. R_1 and C_1 are the pure minimax strategies, but not optimal. The optimal mixed strategies are found to be: $p_1 = 3/4$ and $q_1 = 3/4$, but these optimal strategies are not in equilibrium.

a)

	Column player C_1	C_2
Row player R_1	2,2	3,4
R_2	4,3	1,1

b)



c) d)

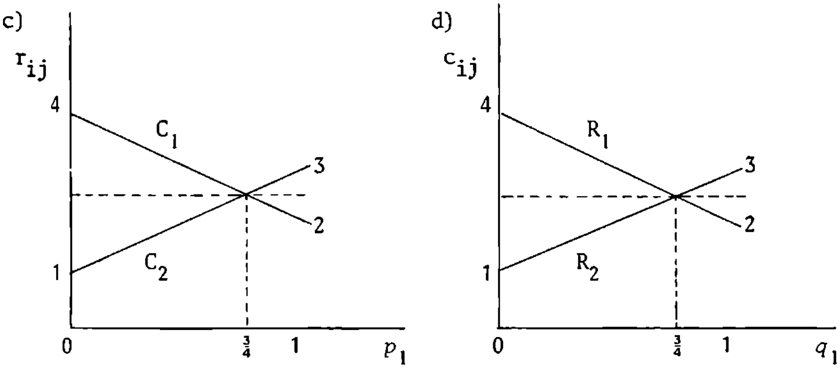


Figure 1.11 Battle of the sexes

(R_2, C_1) and (R_1, C_2) are both equilibrium. The difference with *The Game of the Chicken* is, that in the latter game the non-optimal payoff (2) of these equilibria is less favourable than the non-optimal payoff in the equilibria of *The Battle of the Sexes* (3). This explains where the name *"Mild Chicken"* comes from.

Also for *The Battle of the Sexes* mixed equilibrium strategies can be computed. They are: $p_1^* = \frac{1}{2}$ and $q_1^* = \frac{1}{2}$. The expected gains for these strategies are: $v_r = v_c = 2.5$. It appears that for both equilibria, (R_1, C_2) and (R_2, C_1), the payoff for one player is larger than his expected gains for the mixed equilibrium strategies, while the payoff for the other player is smaller than his expected gains for the mixed equilibrium strategies.

The last example of a two-person nonzero-sum game that will be discussed here, is the game called *"The Hero"* or *"Let George do it"*. There is just a subtle difference with *The Battle of the Sexes*: in *Hero* there are the same equilibria, (R_1, C_2) and (R_2, C_1) as in *The Battle of the Sexes*, but the payoffs are interchanged. The effect of this is, that in *The Hero* an altruistic choice by both players yields (R_2, C_2) and an individualistic choice by both players yields (R_1, C_1), while in *The Battle of the Sexes* the altruistic choice results in (R_1, C_2) and the individualistic choice results in (R_2, C_1) (see Figure 1.12)

In *The Hero* R_1 and C_1 are the pure minimax strategies. The optimal mixed strategies are: $p_1 = \frac{1}{2}$ and $q_1 = \frac{1}{2}$. However, neither is (R_1, C_1) an equilibrium, nor is an equilibrium yielded by the optimal mixed strategies.

The mixed equilibrium strategies are: $p_1^* = 3/4$ and $q_1^* = 3/4$, with expected gains: $v_r = v_c = 2.5$.

1.3. *Games and the study of interpersonal conflicts*

For the game theorist analytical complexities arise in nonzero-sum games, since no unambiguous "solutions" are derived by the normative theory. However, nonzero-sum games are the most interesting games for behavioral researchers to investigate the decision processes in interpersonal conflicts.

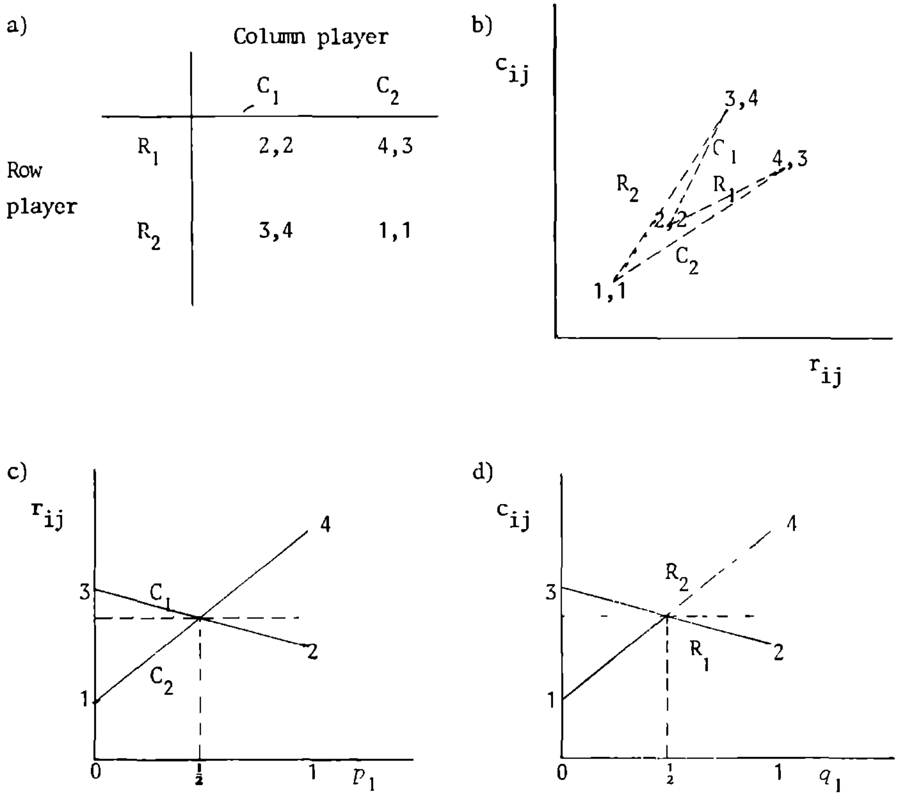


Figure 1.12 The Hero

Psychologists are concerned with determining and interpreting regularities in actual game playing behavior, that is their approach is descriptive rather than prescriptive.

There is a relation between the normative and descriptive theory. The derivations of optimal choices, which are provided by the normative theory, are excellent standards or norms to assess the "rationality" of subjects' choices in experimental games. Comparing actual choices in experimental games with those prescribed by the mathematical normative theory shows that human subjects not always behave in conformity with the normative theory (Krivohlavý, 1974; Lieberman, 1960). For the class of zero-sum games this was demonstrated by, for example, studies of Flood (1952), Vinacke and Arkoff (1957), Atkinson and Suppes (1958, 1959) and Estes (1962).

Even more studies are known on nonzero-sum games, which show great discrepancies between observed behavior and the prescriptions of the normative theory. An excellent overview of the literature can be found in Krivohlavý (1974). In particular, paradigms of games have been studied which leave the normative theory with puzzles and paradoxes, such as the Prisoner's Dilemma Game. Remember that in the PDG if both players choose their dominated strategy, the resulting outcome is jointly better than if both players choose their dominating strategy. For this reason the first strategy of the game matrix in Figure 1.7, R_1 and C_1 respectively, is usually labeled c (for cooperation). If a player in the PDG selects his second strategy, while his opponent selects his first strategy (c), the former player gets his best (most preferred) payoff at the cost of the latter player's payoff. Therefore, the second strategy in the PDG is often labeled d (for defection). In what follows this terminology will be used. Thus, the game matrix of the Prisoner's Dilemma Game is given as

	c	d
c	3,3	1,4
d	4,1	2,2

Research on the Prisoner's Dilemma Game has shown that there are large individual differences in performance which are related to personality characteristics of the players, the payoff matrix, the strategy of the other player, the presence or absence of communication between the players etcetera. See for example Krivohlavý (1974) and Nemeth (1974) for an overview of the relevant literature.

The most frequently used index of behavior (the dependent variable) in studies on the PDG is the relative frequency of choosing c. This frequency is sometimes calculated over a sequence of repeated plays of the game by the same pair of players, and sometimes over several players in a "one shot" experiment or in a grand average over both plays and players (Rapoport and Orwant, 1962). But, as Rapoport and Orwant argue properly, averaging over repeated plays throws away information about the effects of experience; averages over players throw away individual propensities to "cooperate" or to "defect".

It is commonly agreed upon that the choice of an alternative, say *d*, in the PDG on the first play of a sequence of a (great) number of successive plays might have a completely different meaning, in a psychological sense, than the choice of *d* on the last play. Yet choice percentages are utilized quite often as indices of behavior. In studies on PDG the percentage of *c*-choices was used by Scodel, Minas, Ratoosh and Lipetz (1959), Radlow and Weidner (1966), Miller (1967) and Krivohlavý (1974), to mention a few.

McClintock and McNeel (1966, 1967) used the percentage of *d*-choices as main criterion. However, this related to studies on the Maximizing Difference Game (MDG). In the MDG there is no conflict in a game-theoretical sense. For each player there is a dominating strategy, which yields an optimal equilibrium, i.e. optimal for all players, when all players choose their dominating strategy (Figure 1.13).

		Column player	
		c	*d*
Row player	*c*	4,4	2,3
	d	3,2	1,1

Figure 1.13 The Maximizing Difference Game

The psychological conflict involved in MDG is that, although the first strategy is dominating in the sense of maximizing individual and joint gain, the second alternative maximizes the relative gain, i.e. the difference between a player's payoff and that of his opponent. The conflict can be summarized as "give up some gain in order to increase the difference with the opponent's payoff". The risk involved is in $(d\ d)$ · when both players choose so as to maximize the relative gain, they get the least preferred payoff.

In PDG there is a confounding of the motivation to maximize own gain (individualism) and to maximize the relative gain (competitiveness). In MDG these motivations are separated (but now cooperativeness and individualism are confounded). For this reason McClintock and McNeel argue, that in the MDG the percentage of d-choices is a more appropriate index, since the choice of d is less ambiguous in terms of motivation than the choice of c.

We already mentioned that a choice of c (or d) in the PDG may have a different meaning at the beginning of a sequence of successive plays than a c-choice (or d-choice) on the last play. Rapoport (1966b) assigned psychological interpretations to the following outcome sequences:

-(cc)-(cd)-(dd).....unilateral decision to be competitive resulting in suspicion from the other player,
-(dd)-(cd)-(cc).....unilateral offer to cooperate, which is accepted,
-(dd)-(cd)-(dd).....non-accepted offer to cooperate.

A great number of observable and descriptive characteristics of choice sequences are used in the analysis of game playing behavior. Guyer (1968) and Krivohlavý (1967) used the following sequence-dependent parameters.

-$T(cc)_1$ .trialnumber of the first occurrence of a choice of (cc),
-$T(cc)_n$ trialnumber of the first occurrence of n successive choices of (cc),
-$T(dd)_n$:similar criterion as $T(cc)_n$ for (dd),
-$L(cc)$:length of maximum run of successive (cc)-choices,
-$L(dd)$:similar as $L(cc)$ for (dd)-choices,
-$L(cd)$:similar as $L(cc)$ for (cd)-choices.

Other criteria concern observable behavior on the last trials of a number of successive plays (Terhune, 1968, Rapoport and Chammah, 1965, McClintock and McNeel, 1966, 1967), the variance of cooperative choices (Knox and Douglas, 1968), and the skewness of the distribution of c-choices (Rapoport and Chammah, 1965).

To detect interaction effects product-moment correlations related to two players in the same game playing the PDG for n times in succession ($n \geqslant 200$) where computed by Rapoport and Chammah (1965, p. 60) according to the formula

$$\rho = \frac{(cc)(dd) - (cd)(dc)}{\sqrt{(cc + cd)(cc + dc)(dd + cd)(dd + dc)}} ,$$

where (cc) = the frequency of outcomes in the PDG which result when both players choose c,

(dd) = the frequency of outcomes in the PDG which result when both players choose d,

(cd) = the frequency of outcomes in the PDG which result when the first player chooses c and the second player chooses d,

(dc) = the frequency of outcomes in the PDG which result when the first player chooses d and the second player chooses c.

Obviously, this is an ordinary phi-correlation for 2*2-contingency tables:

$\rho = -1$ when all outcomes are (cd) or (dc),

$\rho = +1$ when all outcomes are (cc) or (dd).

Several correlation coefficients were computed:

ρ_0 : the product-moment correlation between both players' choices on the same play;

ρ_1 : the same as ρ_0 except that one player's choice is matched with the other player's immediately preceding choice;

ρ_i $(i = 2,3,4,5,6)$: the same as ρ_0 except that one player's choice is matched with the choice made by the other player i plays previously.

According to Rapoport and Chammah ρ is a measure of the type and strength of influence one player has on the other in repeated plays. Rapoport and Chammah report strong positive correlations in repeated plays of the PDG. In their experiments with the Prisoner's Dilemma Game they had different experimental conditions depending on the way seven different payoff matrices of the PDG-type were administered.

In the *Pure Matrix Condition* ten pairs of players played the same game (that is the same payoff matrix) for three hundred plays, seventy pairs in all or twenty-one thousand outcomes. In this condition the game matrix was displayed to the players throughout the experiment.

In the *Block Matrix Condition* each pair of players played all seven games. A Latin Square Design (Winer, 1971) was used in order to distinguish the effects of the payoff matrix from the effects of learning. There were two Latin squares with each square containing seven orders of the seven games, seven orders of blocks of fifty plays of each game of the seven. There were five pairs of players playing the seven games in a given order, seventy pairs in all. Each pair gave 350 outcomes.

In the *Mixed Matrix Condition* the seven games were presented in random order to each of ten pairs of players, who played seven hundred times or one hunderd plays per game. Here the game matrices were displayed as well.

The *Pure No Matrix Condition* was similar to the Pure Matrix Condition, except that the game matrices were not displayed but the payoffs to both players were announced.

The *Mixed No Matrix Condition* related to the Mixed Matrix Condition in the same manner as the Pure No Matrix Condition did to the Pure Matrix Condition.

The average values of ρ_0 were computed for the five conditions. As Table 1.1 shows, the average correlation was highest in the Block Matrix Condition (.56), where each pair of players played all seven games in the Latin square design.
The average values of ρ_0 were smallest in both No Matrix Conditions: apparently the interactions are weaker in the absence of the payoff matrices.

Table 1.1

Table of ρ-values from Rapoport and Chamah (1965, p. 62)

Condition	ρ_0	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5	ρ_6
Pure Matrix	.46	.51	.46	.42	.40	.38	.36
Block Matrix	.56	.59	.56	.52	.51	.49	.46
Mixed Matrix	.47	.34	.31	.30	.31	.30	.29
Pure No Matrix	.37	.47	.40	.33	.32	.32	.32
Mixed No Matrix	.34	.22	.22	.23	.21	.21	.20

The correlation ρ_0 is not a proper measure for the play to play interaction between players in the same pair. An interaction effect of this sort is reflected in ρ_1. Similarly, ρ_i (i = 2, ..., 6) will measure the degree of interaction of a player's choice with the choices of the other 2, 3, 4, 5 and 6 plays ago. The values of these coefficients for all the conditions are also shown in Table 1.1.

If the behavior of a player is biased by the behavior of the other player in the same pair, this should be reflected in ρ_1 more strongly than in ρ_0. One can also expect that this bias will wear off, when the number of steps (plays) between a player's choice and the choice under consideration of the other in the same pair is increased. Both expectations were realized in the experiments of Rapoport and Chammah as Table 1.1 shows. The expections in both mixed conditions were explained by Rapoport and Chammah as due to the fact that in the mixed conditions the next play generally involved a different game (randomized order of the games!). The payoff matrices could have played a part in influencing the choice of the game, that is there could have been a confounding between a "payoff-effect" and an interaction effect.

Because the foregoing correlation ρ_i ($i = 0,1, \ldots, 6$) were based on frequencies of "cooperative" and "defecting" choices, Rapoport and Chammah concluded that these coefficients are not very suitable as indicators of personal propensities. Instead they proposed two sets of so called contingent propensities: the response-conditioned propensities and the state-conditioned propensities.

The Response-Conditioned Propensities

ξ_i : the probability that player i ($i = 1,2$) chooses *c* following the other's choice of *c* on the preceding play,

η_i : the probability that player i ($i = 1,2$) chooses *c* following his own choice of *c* on the preceding play,

ζ_i : the probability that player i ($i = 1,2$) chooses *c* following his own choice of *d* on the preceding play,

ω_i : the probability that player i ($i = 1,2$) chooses *c* following the other's choice of *d* on the preceding play.

The following interpretations were given by Rapoport and Chammah to these probabilities:

1. ξ is the propensity to respond cooperatively to the other's cooperative response,
2. η is the propensity to "respond" cooperatively to one's own cooperative choice,
3. $1-\zeta$ is a measure of the persistence in the *d*-choice,
4. $1-\omega$ is the propensity to respond noncooperatively to the other's defecting response.

The State-Conditioned Propensities

x_i : the probability that player i ($i = 1,2$) chooses *c* following a (*cc*)-outcome on the preceding play,

y_i : the probability that player i ($i = 1,2$) chooses *c* following a (*cd*)-outcome if $i = 1$ or following a (*dc*)-outcome if $i = 2$ on the preceding play,

z_i : the probability that player i ($i = 1,2$) chooses *c* following a (*dc*)-outcome if $i = 1$ or following a (*cd*)-outcome if $i = 2$ on the preceding play,

w_i : the probability that player i (i =1,2) chooses *c* following a *(dd)*-outcome on the preceding play.

Other stochastic criterions can be derived from the state-conditioned propensities, which appear to be meaningful indices. M is defined by Rapoport and Chammah (1965) as

$$M = \frac{(1-y_1)(1-z_2)}{y_1 z_2} \quad \text{or} \quad \frac{(1-y_2)(1-z_1)}{y_2 z_1}$$

and is related to the "martyr" runs of unilateral states. M is the ratio of the probability that a unilateral state passes to *(dd)* to the probability that the "defector" starts to cooperate while the "martyr" continues to cooperate.

There is an essential difference between games as decision models and games as they are used in social psychological research. In the former approach games are used to devise and study formal models of rational behavior. In the second approach the game matrix is used as a payoff device in situations of interdependence. This latter orientation has generated a myriad of studies on cooperation and competition in experimental games. In these studies the research methodology consists usually of establishing a relationship between an independent and a dependent variable, the latter being commonly some index of (cooperative) behavior.

Recently, the apparent popularity of gaming research in psychology has been criticized in some reviewing articles (Apfelbaum, 1974; Nemeth, 1974; Pruitt and Kimmel, 1977).

Apfelbaum (1974) pointed out the great diversity of theorizing in gaming studies and the lack of an integrative theoretical framework. Another criticism concerned the inconsistent experimental results. As she put it (p. 103-104): "... *most of the investigations have explored different variables, but there has been little systematic attempt to relate the different dimension to one another, and there has been little effort to reconcile the often contradictory experimental evidence.*"

The lack of theory in experimental gaming was also recognized by Pruitt and Kimmel (1977). In an effort to integrate experimental results in the domain of the Prisoner's Dilemma Game and Prisoner's Dilemma type situations they presented a "goal/expectation" theory for behavior in repeated interaction in Prisoner's Dilemma situations. In this theory continuous cooperation is viewed as the result from long-range thinking in which a goal of establishing and/or maintaining continued mutual cooperation is achieved, and at the same time the expectation that the other will cooperate, is formed.

Unfortunately, the experimental findings which were interpreted by Pruitt and Kimmel in the framework of their goal/expectation theory, entailed only measures of the level of cooperation (such as choice percentages) rather than measures of goals and/or expectations. Sequential characteristics of the behavior were not taken into account.

1.4. *In Search of a Parameterization*

The basic form of the pay-off-matrix of the Prisoner's Dilemma Game is

	c	*d*
c	R_1,R_2	S_1,T_2
d	T_1,S_2	P_1,P_2

with $T_i > R_i > P_i > S_i$ and $S_i + T_i < 2R_i$ $(i = 1,2)$ (Rapoport and Chammah, 1965, p. 34). R stands for "reward", T for "temptation", S for "sucker's payoff" and P for "punishment".

This parameterization of the PDG was used by Rapoport and Chammah to derive some indices from the payoffs to relate the observed behavior to.

Starting from the assumption that the behavior of players in the PDG remains invariant if the payoff matrix of the PDG is subjected to a linear transformation Rapoport and Chammah propose the following ratios (1965, p. 41):

$$r_1 = \frac{R - P}{T - S} \quad \text{and} \quad r_2 = \frac{R - S}{T - S}$$

$$0 < r_1 < 1 \qquad\qquad 0 < r_2 < 1$$

In a later study Rapoport (1967) called r_1 the "Index of Cooperation", K. Komorita (1967) interpreted K as "*the incentive to make the c-choice relative to the d-choice.*" (p. 360).
The parameters R, T, P and S are not linked to any psychological theory of game playing behavior. Therefore, this parameterization is not useful to study the effect of variations in the payoff matrix on the psychological structure of behavior. For instance, changes in the frequency of *c*-choices can be conceived of as a function of eight (!) variables. Apart from the fact that this is far too complex to render in a simple model about the conflict in the game, whether and how variations in the payoffs change the player's conception of the game cannot be concluded.

A special subset of games, devised to investigate the relation between game playing behavior and motives of behavior is formed by the separable games (Hamburger, 1969). Pruitt (1967) speaks of decomposed games.

In a *decomposed* or *separable* game each subject (in the dyad) is given a choice between two options, *c* and *d*, such that each option is an ordered pair of numbers, (x,y), where x denotes the payoff to the subject making the choice, and y denotes the payoff to the other person (Messick and McClintock, 1968, p. 7). In Table 1.2 a PDG with a decomposition is shown. As can be seen in part (b), each player receives two payoffs, one from the column he has chosen and one from the column the other has chosen.
A player's final payoff is calculated as the sum of the payoff he has given himself and the payoff the other player has given him.

In fact any game has an infinite number of decompositions, while to each decomposed game there corresponds precisely one game in normal form. Pruitt (1970) showed that different decompositions of the PDG produced different motives which led to differing patterns of behavior.

Table 1.2. Decomposed Prisoner's Dilemma Game

a) game matrix

		Column player	
		c	*d*
Row player	*c*	6,6	0,9
	d	9,0	3,3

b) decomposition

		Choices	
		c	*d*
Payoffs	Own	3	6
	Other's	3	-3

For the subset of separable PDG's Coombs (1973) has proposed a reparameterization of the PDG, which analyzes the payoffs of the game into components with a clearer psychological interpretation than merely the monetary payoffs in terms as "reward", punishment" etcetera.

In Coombs' reparameterization the dilemma of the PDG is reformulated as an instance of an approach-avoidance conflict. The dilemma is interpreted as a choice paradigm having a risk inducement structure. The choice between *c* and *d* is conceptualized as a problem of risk decision making and risk preference (Coombs and Huang, 1970). A player in the PDG is in conflict between an (added) inducement to defect and his vulnerability to retaliation by his opponent. It is a conflict between greet and fear in the context of the level of the game. (see also Coombs, 1975 and Coombs and Avrunin, 1977).

According to Coombs (1973) the strenght of preference for defecting over cooperation would be a distributive model (Krantz and Tversky, 1971):

$$S(d \geqslant c) = [\phi_1(\Delta b) + \phi_2(2a)]\ \phi_3(b),$$

where $\geqslant$ is a binary preference relation and S, ϕ_1, ϕ_2, ϕ_3 are real valued functions defined on $(d \geqslant c)$, Δb, 2a, and b respectively.

Δb = added inducement to defect,
a = vulnerability to retaliation by the other,
b = the "level" of the game.

The payoffs in the game can be written in terms of a, b, and Δb, which makes Coombs' parameterization suitable to simple and direct experimentation. Of course, Coombs' parameterization cannot claim exclusiveness. It is attractive because the interpretation of the parameters links psychological variables with experimental variables, which can be easily manipulated. Experimental results from this manipulation can be easily interpreted, whereas, as we noted before, experimental manipulation of separate payoffs is very complex and cannot be traced back to psychological effects. In chapter 2 of this thesis a detailed exposition of Coombs' reparameterization of the PDG is given.

1.5. *Dynamic decision making and latent behavior*

In real life decisions occur in sequences, and information available for future decisions is likely to be contingent on the nature and consequences of earlier ones. The study of decision processes in such changing situations is called the study of dynamic decision making (Edwards, 1961).

Attempts have been made to explain individual behavior in experimental games as a learning process (Atkinson and Suppes, 1958, 1959; Burke, 1959; Suppes and Atkinson, 1960). These studies were all directed towards the manifest behavior over a number of successive plays and the course of behavior was analyzed to fit a Markov chain model. Although some discrepancies between the normative theory and actual behavior could be accounted for, this approach has been critized on the ground that the application of learning models to social interaction did not start from any psychological theory about social interaction but rather from the apparent usefulness of these models in other areas (Rosenberg, 1968).

One exception must be mentioned here. Foa and Zacks (1959) have developed a stochastic model of social interaction in a dyad, which was based on an earlier paper by Foa (1958). Based on some concepts from Heider's balance theory, Foa developed a classification of behavior in the dyad as to whether a given perceived behavior does or does not conform to the corresponding norm of a subject and/or the other person. However, this classification was concerned with the behavioral effects of social rewards on compliance with learned norms rather than with experimental games or decision making under uncertainty. The stochastic model of Foa and Zacks was based on two basic assumptions. the classification of behavior and learning processes. In their model 256 (!) different states for the two-actor system were identified, which posed serious measurement problems. This was too unwieldy to suggest an estimation procedure for the parameters of the model or to make possible an experimental test.

In a most interesting study by Ofshe and Ofshe (1970) choice behavior in coalition games (three person) was interpreted as resulting from two weighted (and non-observable) utilities utility for money and utility for equity. Applications to non-cooperative games were discussed as well.

The significance of the Ofshe-and-Ofshe study is that decision making in experimental games is not described as choosing an optimal strategy but rather as a stochastic process with non-observable states representing moral or motivational principles. The specific nature of their model did not provide for the analysis of sequences of behavior. Instead choice proportions were analyzed.

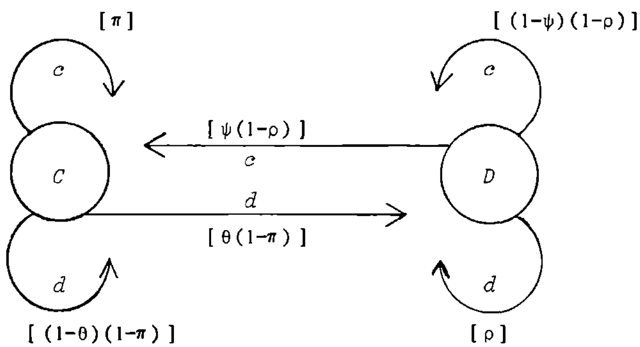
In the studies cited in this section the analysis was directed to the manifest behavior of subjects, such as deriving learning curves, determining asymptotic behavior, comparing choice proportions between different experimental groups or treatments. A major objection to this approach is that characteristics of the players cannot be deduced, unless a posteriori from the course of the manifest behavior.

An original and nicer conceptualization of a two person conflict in the PDG was proposed by Meeker (1971), who started from some hypotheses about the relation between a subject's orientation towards the conflict and his manifest behavior, and provided for changes in this orientation depending on own behavior, other's behavior and the structure of the situation.

In Meeker's theory a distinction is made explicitly between a subject's "*value state*" or social orientation, and his manifest (observable) behavior. *Value states* can be thought of as attitudinal states or dimensions representing a subject's (wittingly or unwittingly) utilized principles of justice. Some behaviors are consistent with subject's orientation, others are not. Changes of subject's orientation are described by a Markov chain model as a function of the consistency of manifest behavior and orientation *(value state)*. To this end use is made of some concepts from consistency theory.

In the mathematical model which was derived from the theory, the probabilities of consistent behavior in C and D are represented by π and ρ respectively. Also, transitions from one state to another after inconsistent behavior are described probabilistically by θ (transition from C to D) and by ψ (transition from D to C).

Meeker's model is shown in the following diagram and matrices, where C and D are two states (C = cooperativeness, altruism and D = rationalism, individualism, competitiveness) and c and d are two behaviors consistent with C and D respectively. Arrows (or loops) indicate state-transitions determined by behavior. The process is assumed to be stochastic.
Probabilities are given in brackets.



The matrices of transition probabilities and response probabilities are:

	C	D
C	$1-\theta(1-\pi)$	$\theta(1-\pi)$
D	$\psi(1-\rho)$	$1-\psi(1-\rho)$

	c	r
C	π	$(1-\pi)$
D	$1-\rho$	ρ

Meeker's approach is not aimed primarily at finding a psychological explanation for subjects' behaving irrationally, but rather to analyze behavior as a function of non-observable *"value states"*. In the next chapter Meeker's model is discussed in detail.

1.6. *Rationality and the theory of two-person justice*

We have already seen that contradictory recommendations arise when the principle of rationality is applied to nonzero-sum games. A different conceptual framework must be designed to reconcile individual and collective rationality. Rapoport and Orwant (1962) suggest that this can be achieved by shifting from a "*rigorously developed normative theory to an experimentally conceived descriptive one*". Unfortunately they (Rapoport and Orwant) did not work out this suggestion.

Two concepts of optimality for two-person, nonzero-sum games prescribe both players in the PDG to choose strategy *d*. One concept arises from the *sure-thing principle*. A strategy or decision satisfies the sure-thing principle, if, no matter what your opponent does, you are at least as well off, and possibly better off, with this strategy in comparison to any other available to you. Thus, according to the sure-thing principle one should choose *d*.

Although many games do not have dominating strategies (which satisfy the sure-thing principle), every finite [1] two-person, nonzero-sum game does have at least one *equilibrium point*, the second concept of optimality, which was introduced by Nash (1950). An equilibrium point is a set of strategies, one for each player, with the property that if all players except one (in the n-person case with $n \geqslant 2$) choose these strategies, the remaining player cannot obtain an outcome preferred to the equilibrium. As is already shown, for the PDG the equilibrium point is *(dd)*, the same result as obtained by application of the sure-thing principle.

As Suppes (1966) points out, the various game-theoretical principles of behavior, that is the sure-thing principle and the equilibrium principle, are aimed at satisfying intuitive ideas of *prudential* rather than *moral* behavior (that is prudential in the sense of acting in one's own best interest without direct concern for others).

This is illustrated as follows. Let a *grading principle* with respect to a social decision situation be a strict partial ordering of outcomes. This definition permits the introduction of a *principle of justice* in addition to principles as minimax or maximizing expected utility. For example, let J_i (referring to player i with i = 1,2) be the grading principle called *more just than* in the Prisoner's Dilemma, which yields the following diagram:

$$\begin{tikzcd}
(cc) \arrow[d] & (dc) & (cd) \\
(dd) & &
\end{tikzcd}$$

The arrow running from *(cc)* to *(dd)* means: "*(cc)* is preferred to *(dd)*". Clearly, $J_1 = J_2$ (and J_i is asymmetric and transitive).

Since J_i relates neither *(dc)* nor *(cd)* to any of the other outcomes, it seems as if this principle is idle. However, it can be shown

1) A game is called finite if each player has a finite number of strategies.

that this grading principle may lead to a moral principle by introducing some further concepts.

If outcomes which are not dominated under the relation J_1 by any other outcome, are designated as (J_1)-*admissible outcomes*, it is seen that (*cc*), (*dc*) and (*cd*) are (J_i)-admissible. Then, let a *point of justice* be any set of strategies, one for each player, such that choosing these strategies yields an admissible outcome. Now, several justice-oriented rules of behavior are suggested by Suppes (1966, p. 302):

I. *If $J_1 = J_2$ (in a two-person situation) and there is a unique point of justice, the strategy belonging to this point ought to be chosen.*

Since there is no unique point of justice in the Prisoner's Dilemma, rule (I) is not applicable.

Next, let a *justice-saturated strategy* (with respect to J_i) for player i be a strategy such that whatever strategies are chosen by the other players the resulting set of strategies is a (Ji) point of justice. The ethical rule of behavior is then:

II. *If for any player the set of justice-saturated strategies is non-empty, he ought to choose one.*

In the Prisoner's Dilemma each player has a unique justice-saturated strategy, namely *c*. If both players jointly choose this strategy, (*cc*) is attained. However, as is immediately clear, the "ethical" player using rule (II) is at a definite disadvantage against a "prudential" player, which renders Suppes' theory of justice of very little use.

The principle weakness of Suppes' theory is, that it does not provide for prediction or anticipation of the strategy-choices by the other player(s). Because of this shortcoming his ethical rules of behavior must be interpreted as unconditional imperatives, and as such his theory is just another normative theory.

Another structural analysis of the PDG is presented by Burns and

Meeker (1973, 1974). Their theory is characterized by (i) descriptive rather than normative or prescriptive models of social behavior; (ii) emphasis on multidimensional processes and on structural relationships rather than on unidimensional quantities (such as 'utility'), (iii) the thesis that evaluation, decision making, and interaction processes cannot be understood apart from the social context in which they occur.

Burns and Meeker assume, that the actors (players) in a social decision situation (such as the PDG) evaluate and rank a finite set of outcomes which can be represented as points (vectors) in some n-dimensional space $\mathbb{R}^n$. It is further assumed that the various dimensions describing the outcomes may have different weights (importance) for any player and also that these weights may differ over players.

Different order relations on the outcomes may be established by different evaluation procedures (e.g. maximization, optimization, lexicographic ordering, functional ordering, weak component ordering).

Choice behavior consists of selecting an alternative or subset of alternatives from among a set of alternatives (Burns and Meeker, 1973, p. 147). The bases of choice are several: selection of the most preferred alternative evaluated by the actor on the basis of goals and values, selection prescribed by authority or tradition, or random selection.

Given a preference structure over outcomes different decision procedures may be used to construct a preference structure over alternatives.

In Burns and Meeker's view outcome evaluation and choice behavior may depend not only on an actor's own personal preferences but on those of other actors with whom he interacts or has social relationships as well (1974, p. 38). They assume that actors' social attitudes or orientations with respect to self or other actors are metaprocesses which act upon preference structures and decision procedures of the actors involved. These metaprocesses may yield modifications of preference structures. Changes in preference structure may affect interaction patterns in game playing behavior. Specifically, it is demonstrated by Burns and Meeker that the Prisoner's Dilemma is "resolved" as a result of changes in outcome-preferences through actors' shift from *self-orientation* to *other-orientation*.

Although Burns and Meeker have developed a theory of behavior, they have not reformulated or revised the concept of rationality. Rather they say that payoffs in any formally presented game, do not represent the players' utilities, and that utilities may change during the course of interaction. However, if preferences over outcomes are modified so as to change the very nature of the game, a resulting resolution of the game should be dismissed for being improper.

A very original approach to the theory of rational behavior is developed by Nigel Howard (1966, 1971) in his theory of meta-games. It is a descriptive theory of actual behavior which is based on a broadening of the concept of rationality by passing from the space of strategies that define a game in normal form to a space of metastrategies, defined as the strategies that would be available to a player if he knew the strategy choices of the other player(s) (Rapoport, 1974, p. 13). The theory of metagames is a formalized way of conveying intent as well as immediate behavior (Shubik, 1970, p. 190), which singles out stable outcomes in the game-theoretic sense which are no "solutions" of the game in classical game-theoretic definitions. We feel that in his theory Howard has succeeded in formalizing Suppes' principles of prudential and moral behavior. There, both principles are principles of (meta)rational behavior in the sence of optimizing behavior.

1.7. *General outline*

In this thesis a theory of two person interaction for the Prisoner's Dilemma Game is employed, which explains the dynamic process of decision making between cooperation and competition through the concept of *value states* as unobservable states which represent the values or norms the decision maker adheres to. The role of these *value states* and the behavior of the decision makers will be described by a latent Markov chain model.

Chapter 2 treats the *value state* model as it was formulated originally by Meeker (1971). An experiment is described in this chapter, which investigates the psychological nature of *value states* by relating the parameters of a latent Markov chain model to the parameters of Coombs' recent parameterization of the PDG in which choices in the PDG are interpreted in terms of Coombs' theory of risk and risk preference. The model, though, is only applied to the restricted condition of a benevolent, that is 100% cooperative, opponent.

In chapter 3 the model of chapter 2 is reformulated to make it applicable to situations of uncertainty. This generalized model is applied to data from an experiment with three different patterns of behavior of the opponent in a Prisoner's Dilemma Game (i.e. three levels of cooperativeness). The model applies only to situations where the opponent's behavior is unresponsive to subject's own behavior.

A special section in this chapter is devoted to problems relating the estimability of models with non-observable states. It is found that the usual techniques, such as the method of moments, maximum likelihood and least squares estimation, cannot be applied to the value state model. Instead, an iterative procedure is employed.

Expanding the empirical adequacy of the model to the situation where the other can reciprocate and retaliate in the Prisoner's Dilemma Game, that is unrestricted game playing by two real players, requires expanding the theoretical adequacy of the value state model. This is inspired by Nigel Howard's theory of metagames, which is described in chapter 4. Also in chapter 4 the usefulness of some assertions of the theory of metagames, which might be adopted for a general model of dynamic decision making in the Prisoner's Dilemma Game, is investigated in an experiment.

In chapter 5 the empirical and theoretical results of the preceding chapters are put together in a reformulation of the value state theory for the Prisoner's Dilemma Game, which is formalized in a dual latent Markov chain model.

An estimation procedure is proposed which has some characteristics of maximum likelihood estimation and which is labeled (tentatively) as Quasi Maximum likelihood or Incomplete Maximum likelihood

Apfelbaum, E. On conflicts and bargaining. In: L. Berkowitz (Ed.), *Advances in experimental social psychology*, vol. 7. New York: Academic Press, 1974.

Atkinson, R.C. and Suppes, P. An analysis of two-person game situations in terms of statistical learning theory. *Journal of Experimental Psychology*, 1958, *55*, 369-378.

Atkinson, R.C. and Suppes, P. Applications of a Markov model to two-person Noncooperative games. In: R.R. Bush and W.K. Estes (Eds.), *Studies in Mathematical Learning Theory*. Standford, 1959.

Burke, C.J. Applications of a linear model to two-person interactions. In: R.R. Bush and W.K. Estes, *Studies in Mathematical Learning Theory*, Standford, 1959.

Burns, T. and Meeker, L.D. A Mathematical Model of Multi-Dimensional Evaluation, Decision Making, and Social Interaction. In: J.L. Cochrane and M. Zeleny (Eds.), *Multiple Criteria Decision Making*, South Carolina: The University of South Carolina Press, 1973.

Burns, T. and Meeker, L.D. Structural Properties and Resolutions of the Prisoner's Dilemma Game. In: A. Rapoport (Ed.), *Game Theory as a Theory of Conflict Resolution*. Dordrecht-Holland: D. Reidel Publishing Company, 1974.

Coombs, C.H. A Reparameterization of the Prisoner's Dilemma Game. *Behavioral Science*, 1973, *18*, 424-428.

Coombs, C.H. Portfolio theory and the measurement of risk, In: M.F. Kaplan and S. Schwartz (Ed.), *Human Judgment and Decision Processes*. New York (Academic Press), 1975.

Coombs, C.H. and Avrunin, G.S. Single peaked functions and the theory of preference. *Psychological Review*, 1977, *84*, 216-230.

Coombs C.H., Dawes, R.M. and Tversky, A. *Mathematical Psychology. An Elementary Introduction*. Englewood-Cliffs (N.J.): Prentice-Hall, Inc., 1970.

Coombs, C.H. and Huang, L.C. Polynomial psychophysics of risk. *Journal of Mathematical Psychology*, 1970, *7*, 317-338.

Edwards, W. Behavioral Decision Theory. *Annual Review of Psychology*, 1961, *12*, 473-498.

Estes, W.K. Theoretical treatment of differential reward in multiple choice learning and two-person interactions. In: J.H. Griswell et al. (Eds.), *Mathematical Methods in Small Group Processes*. Standford, 1962.

Flood, M.M. Some Experimental Games. Research Memorandum RM-789, The Rand Corporation, Santa Monica, 1952.

Foa, U.G. Behavior, Norms, and Social Rewards in a Dyad. *Behavioral Science*, 1958, *3*, 323-334.

Foa, U.G. and Zacks, S. A Stochastic Facet Theory of Social Interaction in the Dyad. Technical (Scientific) Note No. 1, The Israel Institute of Applied Social Research, 1959.

Gyuer, M. Response-dependent parameter change in the Prisoner's Dilemma Game. *Behavioral Science*, 1968, *13*, 205-219.

Hamburger, H. Separable games. *Behavioral Science*, 1969, *4*, 121-132.

Howard, N. The Theory of Metagames. *General Systems*, 1966, *XI*, 167-186.

Howard, N. *Paradoxes of Rationality: Theory of Metagames and Political Behavior*. Massachusetts: MIT Press, 1971.

Knox, R.E. and Douglas, R. Low pay-offs and marginal comprehension: two possible constraints upon behavior in the Prisoner's Dilemma Game. Report to the W.P.A. Meeting, San Diego, California, 1968.

Komorita, S.S. A Model of the N-Person Dilemma-Type Game. *Journal of experimental Social Psychology*, 1976, *12*, 357-373.

Krantz, D. and Tversky, A. Conjoint-measurement analysis of composition rules in psychology. *Psychological Review*, 1971, *78*, 151-169.

Krivohlavý, J. Grausamkeitzkriterien. *Forschungsbericht* , 1967.

Krivohlavý, J. *Zwischenmenschliche Konflikte und Experimentelle Spiele*. Bern: Verlag Hans Huber, 1974.

Lieberman, B. A failure of game theory to predict human behavior. Memo. SP-101. Laboratory of Social Relations, Harvard University, 1960.

McClintock, C.G. and McNeel, S.P. Reward and score feed-back as determinants of cooperative and competitive behavior. *Journal of Personality and Social Psychology*, 1966, *4*, 606-613.

McClintock, C.G. and McNeel, S.P. Prior dyadic experience and monetary reward as determinants of cooperative and competitive game behavior. *Journal of Personality and Social Psychology*, 1967, *5*, 282-294.

Meeker, B.F. Value Conflict in Social Exchange: a Markov Model. *Journal of Mathematical Psychology*, 1971, *8*, 389-403.

Messick, D.M. and McClintock, C.G. Motivational Bases of Choice in Experimental Games. *Journal of Experimental Social Psychology*, 1968, *4*, 1-25.

Miller, R.R. No play: a means of conflict resolution. *Journal of Personality and Social Psychology*, 1967, *6*, 150-156.

Nash, J.F. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, U.S.A., 1950, *36*, 48-49.

Nemeth, C. A critical analysis of research utilizing the Prisoner's Dilemma paradigm for the study of bargaining. In: L. Berkowitz (Ed.), *Advances in Experimental Social Psychology*, vol. 7. New York: Academic Press, 1974.

Ofshe, L. and Ofshe, R. *Utility and Choice in Social Interaction*. Englewood Cliffs (N.J.): Prentice-Hall Inc., 1970.

Pruitt, D.G. Reward structure and cooperation: The decomposed Prisoner's Dilemma Game. *Journal of Personality and Social Psychology*, 1967, *7*, 21-27.

Pruitt, D.G. Motivational processes in the decomposed Prisoner's Dilemma Game. *Journal of Personality and Social Psychology*, 1970, *14*, 227-238.

Pruitt, D.G. and Kimmel, M.J. Twenty years of experimental gaming: critique, synthesis, and suggestions for the future. *Annual review of Psychology*, 1977, *28*, 363-392.

Radlow, R. and Weidner, M.F. Unenforced commitments in cooperative and non-cooperative nonconstant-sum games. *Journal of Conflict Resolution*, 1966, *10*, 497-505.

Rapoport, A. *Two-person Game Theory*. Ann Arbor: The University of Michigan Press, 1966. (a).

Rapoport, A. Games as tools of psychological research. M.H.R.I. Report, 1966. (b).

Rapoport, A. A note on the "index of cooperation" for Prisoner's Dilemma. *Journal of Conflict Resolution,* 1967, *11*, 101-103.

Rapoport, A. *N-person Game Theory.* Ann Arbor: The University of Michigan Press, 1970.

Rapoport, A. Prisoner's Dilemma - Recollections and Observations. In: A. Rapoport (Ed.), *Game Theory as a Theory of Conflict Resolution.* Dordrecht-Holland: D. Reidel Publishing Company, 1974.

Rapoport, A. and Chammah, A. *Prisoner's Dilemma.* Ann Arbor: University of Michigan Press, 1965.

Rapoport, A. and Orwant, C. Experimental games: a review. *Behavioral Science,* 1962, *7*, 1-37.

Rosenberg, S. Mathematical Models of Social Behavior. In: Lindzey and Aronson (Eds.), *Handbook of Social Psychology,* Vol. I, 1968.

Scodel, A., Minas, J.S., Ratoosh, P. and Lipetz, M. Some descriptive aspects of two-perso nonzero-sum games. *Journal of Conflict Resolution,* 1959, *3*, 114-119.

Shubik, M. Game theory, behavior, and the paradox of the Prisoner's Dilemma: three solutions. *Journal of Conflict Resolutions,* 1970, *2*, 181-193.

Suppes, P. Some Formal Models of Grading Principles. *Synthese,* 1966, *16*, 284-306.

Suppes, P. and Atkinson, R.C. *Markov Learning Models for Multiperson Interaction.* Standford: The Standford University Press, 1960.

Terhune, K.W. Motives, situation, and interpersonal conflict within Prisoner's Dilemma. *Journal of Experimental and Social Psychology,* 1970, *6*, 187-204.

Vinacke, W.E. and Arkoff, A. An Experimental study of coalitions in the triads. *American Sociological Review,* 1957, *22*, 1-12.

Von Neumann, J. and Morgenstern, O. *Theory of Games and Economic Behavior.* Princeton University Press, 1944.

Winer, B.J. *Statistical Principles in Experimental Design,* 2nd ed. New York: McGraw-Hill, 1975.

Chapter 2: Latent Markov chain analysis of a value conflict in Prisoner's Dilemma Games

A.L. van der Sanden, 'Latent Markov chain analysis of a value conflict in Prisoner's Dilemma Games', *Br. J. math. statist. Psychol.*, 1978, *31*, 126-143.

If a subject's opponent in iterated plays of the Prisoner's Dilemma Game is non-responsive and 100 per cent cooperative, the subject is in conflict between reciprocating the cooperative acts of the other and choosing the 'rational' alternative which maximizes his payoff at the cost of the other.
It is assumed that this is a conflict between opposed values which a subject can hold. The process of decision-making in this value conflict is described in a Markov model, which relates non-observable value states to the observable behaviour. The relationships between the parameters of the model and a new parameterization of the Prisoner's Dilemma Game are examined to add more content to the constructs of value states and to gain more insight of the dynamics of the decision-making process in the game.

2.1. *Introduction*

In the study of human game behaviour it is often assumed, implicitly or explicitly, that the rank order of preferences for the outcomes is different for different subjects (players). It is said then that the interests of the player are in conflict. Psychologically, the most interesting games are those in which the interests of the players are partly coincident and partly opposed, because then one can postulate not only a conflict among the players but also inner conflicts within the players (Rapoport & Chammah, 1965). In two-person non-zero-sum games some outcomes are jointly better for both players than other outcomes. In such games the subject is in conflict between pursuing the common interests or striving after his own individual interests.

Discussing the well-known Prisoner's Dilemma Game (PDG), Luce & Raiffa (1957, p.96) argued that for a single play of the game the only rational and completely justified choice for both players is that alternative which dominates the other alternative. In the example shown in Table 2.1 alternative *d* is dominant over *c* for both players.

Table 2.1 Example of Prisoner's Dilemma Game

		Player Y	
		c	*d*
Player X	*c*	2,2	-3,4
	d	4,-3	-1,-1

Thus each player's rational choice is *d*, but the resulting outcome, (*dd*), is less preferrt than (*cc*). This constitutes the dilemma.

With multiple iterations it is possible for each payer to take into account the past behaviour of the other when making a new decision. A clear demonstration of such an interaction effect is the 'lock-in' effect found by Rapoport abd Chammah (1965).

In the past decade several authors have explained human behaviour in experimental games by postulating motivational tendencies or orientations as determinants of subjects' choices in iterative plays (Messick & McClintock, 1968; Pruitt, 1970; Wyer,1971; Griesinger & Livingston, 1973; Friedland *et al.*, 1974; Kuhlman & Marshello, 1975). A subject may have different motivational orientations in playing a experimental game. An individual may try to maximize his own gain (individualism), try to maximize the difference between his own gain and his opponent's gain (competition), or try to maximize the joint gain (cooperation). When a subject is motivated by more than one value simultaneously and these values give him incompatible prescriptions, he may find himself in a conflict, a value conflict.

2.2. *Meeker's conflict model*

Meeker (1971) has developed a Markov model to describe the process of decision-making by a subject who must make a series of choices between reciprocating the cooperative acts of another person and maximizing his own pay-off in a situation in which the other cannot retaliate. The model is based on the following hypothetical situation: S (the subject) and O (the other) are engaged in a social exchange of the PDG type. O has been cooperative in the past and O's future behaviour is not contingent on S's present behaviour. S, who expects O to be cooperative on the next play, must choose between reciprocating the cooperative acts of O and being purely 'rational'. Since S need not fear retaliation the conflict is purely intrapersonal.

The two strategies open to S are each consistent with a different set of motivations. These mutually exclusive sets are called 'value states'. The value state state which is characterized by altruism and reciprocity is referred to as the 'reciprocating state' (C). The other value state, which is characterized by individualism, competition and rationalism, is called the 'rational state' (D).

The following axioms about the properties of value states are adopted from Meeker (1971).

Axiom 2-1: There are two states, reciprocating (C) and rational (D). There are two responses, c (cooperate) and d (defect). On any trial n (n = 1,2,...) S is in exactly one state and gives one response; 1.e.

$$P[D_n] = 1 - P[C_n], \tag{2.1}$$

$$P[d_n] = 1 - P[c_n]. \tag{2.2}$$

In each value state one act is defined as the most desirable or consistent with the values which are included in that state. It is assumed that in each value state the consistent behaviour is chosen with probability greater than 0.50 on each trial.
Axiom 2-2 defines the probabilities for consistent behaviour.

Axiom 2-2:

$$P[c_n | C_n] = \pi \quad (\pi > 0.50), \tag{2.3}$$

$$P[d_n | D_n] = \rho \quad (\rho > 0.50). \tag{2.4}$$

After each response S may change his value state on the next trial. Meeker adapted some ideas from Heider's balance theory and Festinger's theory of cognitive dissonance which are incorporated in his theory as transition axioms. It is assumed that as soon as S behaves inconsistently, that is, choosing d when occupying state C or choosing c when occupying state D, he experiences discomfort and will tend to adjust by changing his value state with a certain probability.

There may be a number of reasons why S does not always change his value state. According to Meeker there is '... some psychological cost involved in changing a value state (admitting he was wrong)'. As long as S behaves consistently he will not change his value state.

Axiom 2-3:

$$P[C_n | C_{n-1}, c_{n-1}] = 1, \tag{2.5}$$

$$P[D_n | C_{n-1}, d_{n-1}] = \theta. \tag{2.6}$$

Axiom 2-4:

$$P[D_n | D_{n-1}, d_{n-1}] = 1, \tag{2.7}$$

$$P[C_n | D_{n-1}, c_{n-1}] = \psi. \tag{2.8}$$

These axioms define a latent Markov chain with four states. Each state is a combination of a value state and a response. Table 2.2 gives the transition probabilities for the Markov chain. The latent process and the axioms can be rendered in a tree diagram as in Fig.2.1.

In an experiment Meeker estimated the value of π, ρ, θ and ψ in three different conditions of conflict between rationality and reciprocity. In two conditions the game played was a real PDG and the third game was a 'degenerated' PDG, in that the subject's decision had no effect on his own score (see Table 2.3). In the first condition

Table 2.2 Transition matrix of the latent Markov chain

Trial n / Trail n - 1	$<C,c>$	$<C,d>$	$<D,c>$	$<D,d>$
$<C,c>$	π	$(1-\pi)$	0	0
$<C,d>$	$(1-\theta)\pi$	$(1-\theta)(1-\pi)$	$\theta(1-\rho)$	$\theta\rho$
$<D,c>$	$\psi\pi$	$\psi(1-\pi)$	$(1-\psi)(1-\rho)$	$(1-\psi)\rho$
$<D,d>$	0	0	$(1-\rho)$	ρ

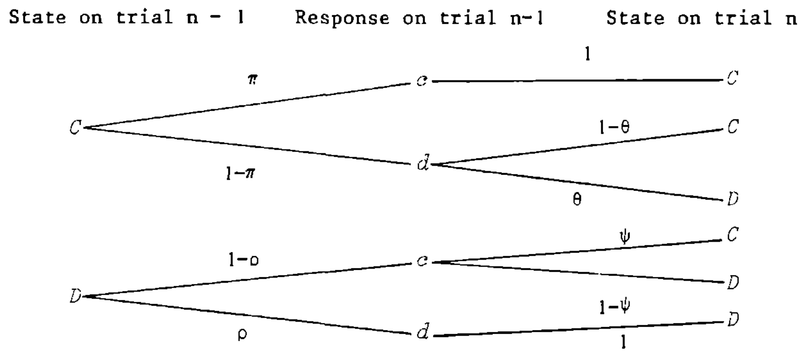


Figure 2.1. Tree diagram showing transition and response probabilities.

the choice of c was labelled "For the group" and the choice of d was labelled "For myself". In the second condition there was no group orientation; otherwise there were no differences between conditions 1 and 2. In the third condition S had no influence on his own pay-off but he could help O or not help O.

The experiment was run over 50 trials. One subject, S, was told on each trial what the other, O, had chosen, while O was never informed about S's choices. As a matter of fact, the feedback S received on each trial was prearranged, and S was always informed that O had been cooperative (i.e. had chosen alternative c).

Since O was never informed about S's behaviour, O's future behaviour was not contingent on S's present behaviour.

Table 2.3 Games from Meeker's experiment; game (a) is a real PDG and game (b) is a PDG with no dominating strategy for S.

(a) Conditions 1 and 2

		O: c	O: d
S	c	2, 2	-3, 3
	d	3, -3	-2, -2

(b) Condition 3

		O: c	O: d
S	c	2,2	0,2
	d	2,0	0,0

At each trial the subjects had to choose a course of action and guess what the other person's choice was going to be. From these guesses one could deduce whether the subjects had learned what the feedback pattern was. Meeker found that by trial 10, and usually at an earlier trial, all subjects had begun to guess c consistently. Therefore the data, consisting of c and d choices, from the last 40 trials were supposed to fulfil the conditions of the theory and were used for the estimation of the four parameters, π, ρ, θ and ψ.

The estimates obtained from Meeker's experiment are presented in Table 2.4. Because the parameters were not estimated by a least-squares method or a maximum-likelihood method,but by an iterative process, no measure of the goodness of the solution found was available. As a partial check Meeker compared some probabilities predicted by the estimated parameters with the corresponding proportions observed in the data. None of these proportions was used in the estimation procedure. The largest difference for all conditions was 0.04.

Table 2.4 Estimates of the parameters after iteration procedure (from Meeker, 1971, p. 398)

Condition	π	ρ	θ	ψ	P[C]
1	0.89	0.69	0.67	0.38	0.62
2	0.85	0.92	0.47	0.69	0.44
3	0.91	0.74	0.29	0.41	0.80

NB. P[C] is the probability that a subject is in value state C.

In formulating his theory Meeker defined the nature of a value state somewhat vaguely. The only assumptions made about its properties are that it influences behaviour and that it may change after inconsistent behaviour. With the aid of experimental results one may give more content to this hypothetical construct. The parameters π and ρ (the probabilities of consistent behaviour) indicate the extent to which a value state influences behaviour.

Meeker noted that the value of ρ was highest in condition 2 in which there was no group orientation; in the other conditions there was either a reduction in cost or no cost attached to choosing c. The value of π was similar for all conditions. According to Meeker, while the influence of state D on behaviour alters with changes in the cost of the behaviour and with the experimental instructions, a property of state C is that its influence on behaviour is not affected by these changes. He hypothesized that state C is primarily oriented to the behaviour of the other person and that since this behaviour was the same for all conditions, the value of π did not vary. He further argued that π should be more sensitive than ρ to variations in the behaviour of the other subject.

The parameters θ and ψ indicate the extent to which a value state is influenced by inconsistent behaviour. Meeker interpreted a high value of θ, the probability of changing from C to D, as an indication of high conflict between a d response and state C. From Table 2.1 it follows that the d response was most in conflict with state C when there was a cost to the other person attached to d and when there is a group orientation.

The highest value of ψ, the probability of changing from D to C, is found in condition 2 in which there was a cost to the subject attached to c and there was no group orientation.

With reference to Meeker's interpretation of the parameters ρ and ψ, inspection of Table 2.4 suggests that ρ and ψ could be correlated. Nevertheless, Meeker concluded that the total analysis of the four parameters indicates that "*...perhaps the effects of values on behaviour and the effect of behaviour on value are independent of each other*" (1971, p. 400).

A number of questions about Meeker's theory remain unanswered. One may wonder whether the conflicts in all three conditions are substantially the same. Meeker assumes implicitly that they are. Yet it is immediately clear that in condition 3 the intrapersonal conflict is essentially different from those in the other conditions. In condition 3 there is no question of individualism in value state D, because S could not profit from choice d. The orientation of the situation and the structure of the game being equal, what is the impact of varying pay-offs and of varying behaviour of O on the values of Meeker's parameters? Answers to these questions may add more psychological meaning to the parameters of the mathematical model and to the construct of a value state. To vary the pay-offs of a PDG in a systematic way we need a useful parameterization of the game. By a useful parameterization we mean that there is a simple and direct correspondence between the experimental variables and some psychological concepts.

2.3. *Coombs' reparameterization of the PDG*

The general properties of the pay-off matrix for a PDG are formulated by Scodel *et al.* (1959). For both players the pay-offs for the cooperative choices (alternative c) and the competitive choices (alternative d) must satisfy the rules shown in Table 2.5. An analogous formulation can be found by Rapoport and Chammah (1965, pp. 33-36).

Table 2.5 Rules for the pay-off matrix of a PDG

I. $2X_1 > X_2+X_3 > 2X_4$

II. $X_3 > X_1$

III. $X_3 > X_2$

IV. $X_4 > X_2$

		Player Y	
		c	*d*
Player X	*c*	X_1, X_1	X_2, X_3
	d	X_3, X_2	X_4, X_4

This rather simple parameterization is confined to those forms of the game which are symmetric for the players. However, the most severe restriction of this parameterization is that it is totally devoid of psychological meaning;that is, it is not related to some psychological theory of game-playing behaviour.

A special subset of PDG's, which allows for asymmetric games, is formed by the so-called "separable games". These are games in which each pay-off to each player can be expressed as the sum of two contributions, one from each player depending on his choice. It is the subset of separable PDGs for which Coombs (1973) has proposed a reparameterization, which is inspired by risk theory, and which relates the independent variables to some theoretical variables in a simple and direct way. Coombs' reparameterization is as follows. Let :

$g = (y,z)$ indicate a gamble with two mutually exclusive outcomes, y and z with $y \geqslant z$;

$g = (0,0)$ be the basic gamble from which others will be constructed, it is interpreted as the status quo;

$a(g) = (a, -a)$ be a transformation in which $y = a$ and $z = a$ and where $a \geqslant 0$.

$b(g) = (b,b)$ be a transformation in which an amount b is added to both outcomes y and z.

The two transformations applied together yield the gamble

$$g_1 = (a + b,\ b - a) \tag{2.9a}$$

Let Δb be an increment to b. This yields the gamble

$$g_2 = (a + b + \Delta b,\ b - a + \Delta b). \tag{2.9b}$$

The gambles correspond to the two available alternatives in a PDG. If we add the subscripts s (for subject) and o (for other) to the parameters, the pay-off structure of PDG can be formulated as is indicated in Table 2.6. If the values for the parameters are set to $a_s = a_o = 2.5$, $b_s = b_o = -0.5$, and $\Delta b_s = \Delta b_o = 2$, then we get the game shown in Table 2.1.

Table 2.6 Coombs' parameterization of the Prisoner's Dilemma Game

		The Other (O) *c*		The Other (O) *d*	
Subject (S)	*c*	a_s+b_s,	a_o+b_o	b_s-a_s,	$a_o+b_o+\Delta b_o$
	d	$a_s+b_s+\Delta b_s$,	b_o-a_o	$b_s-a_s+\Delta b_s$,	$b_o-a_o+\Delta b_o$

In an asymmetric game the values of the parameters are different for both players. Although this yields six parameters, three for each subject, only three distinct psychological variables are involved, because the parameters play symmetric roles for the subjects (Coombs, 1973, p. 425). The gamble g_2 is a translation of g_1 by an amount Δb. This parameter is designated by Coombs as the added inducement to the subject to defect. The parameter b is an additive constant for all outcomes. It sets the 'level' of the game. Suppose the other behaves randomly, then b is the expected pay-off if the subject constantly chooses the first alternative (plays g_1). There may be a specific relationship between b and Δb in that the effect of Δb decreases as the value of b increases in line with the theory of marginal decreasing utility.

The interpretation of the parameter a is a little bit more subtle. In the context of individual risky decision making it is a risk parameter; in the gambles g_1 and g_2 the value $2a$ is the difference between the possible outcomes of the gambles. The larger the value of a, the greater is the risk involved in both gambles. The pay-off, which a subject will get after he has chosen an alternative (g_1 or g_2), depends upon the other's choice. The control that Other has over Subject's outcomes is exactly $2a$. The larger the value of a the more vulnerable a subject is to retaliation by the other, after he (the subject) has defected on an earlier trial (made the d response). In this respect the parameter a represents the risk of the game.

Finally, it should be kept in mind that the parameters for both players have identical roles; Subject also has control over Other's outcomes. As with b and Δb there may be an interaction effect between b and a such that the effect of a is decreased as b increases.

2.4. *Hypotheses*

Coombs' reparameterization of the PDG enables us to vary the pay-off structure of this game in a systematic way such that changes in the outcomes can be interpreted psychologically. Is is possible now to hypothesize about the effects of varying pay-offs of the PDG on the values of the parameters of Meeker's conflict model. In what follows, we suppose that $b_s = b_o$.

Meeker found that unless there was a special orientation towards the game then π and ρ, the probabilities of consistent behaviour in the two value states, remained fairly constant over the conditions he used. Therefore, it is assumed that the response probabilities in the *latent states are invariant to changes in the pay-off matrix.*

The parameters θ and ψ, probabilities of changing value state after inconsistent behaviour, are measures of the effect of behaviour on the value state. Meeker's theory does not give any specific prediction about the relation between the pay-off structure of the game and the change of state parameters.

The reciprocating state is defined as 'directed towards the other'. The d response produces a cost to Other. This cost, expressed by a_o, conflicts with the reciprocating state. The variable a_s, expressing the control O has over S's outcome, can be taken as a measure of the 'cooperativeness' of O, provided O is always cooperative. The more cooperative O is, the more conflicting is a d response of S in state C. Thus, increasing a_s and/or a_o raises the perceived cost of the conflict thereby increasing the probability (θ) that S will move into the rational state.

In the rational state S is self-interested. The larger the value of Δb_s, the larger the cost to S which is attached to the c response and the greater the 'irrationality' of S's behaviour. Thus, it is supposed that increasing Δb_s will increase the probability (ψ) of going to the reciprocal state after inconsistent behaviour in state D.

Increasing Δb_o increases the 'irrationality' of O's cooperative behaviour. It is supposed that increasing Δb_o weakens the conflict by a C response in state D, i.e. increasing Δb_o will decrease ψ.

2.5. *Experimental procedure*

An experiment was run to test the above-mentioned hypotheses. Two levels were used for each of the experimental parameters a_s, Δb_s, a_o and Δb_o. The levels are indicated by the superscripts + and -. The values b_s and b_o were held constant and equal throughout the experiment. The values for the parameters are listed in Table 2.7. The design was

Table 2.7 Values of the experimental parameters

Parameters	Levels +	Levels -
a_s	2·5	1·5
Δb_s	2·0	1·0
a_o	2·5	1·5
Δb_o	2·0	1·0

$b_s = b_o = -0{\cdot}5$.

Table 2.8 Factorial scheme showing the eight treatment combinations selected; the numbers in the cells index the experimental conditions used in the present study.

		a_o^+		a_o^-	
		Δb_o^+	Δb_o^-	Δb_o^+	Δb_o^+
a_s^+	Δb_s^+	1			2
	Δb_s^-		3	4	
a_s^-	Δb_s^+		5	6	
	Δb_s^-	7			8

Table 2.9 The resulting PDGs for the eight treatment combinations†

(a) Condition 1

S \ O	*c*	*d*
c	2, 2	-3, 4
d	4, -3	-1, -1

(b) Condition 2

S \ O	*c*	*d*
c	2, 1	-3, 2
d	4, -2	-1, -1

(c) Condition 3

S \ O	*c*	*d*
c	2, 2	-3, 3
d	3, -3	-2, -2

(d) Condition 4

S \ O	*c*	*d*
c	2, 1	-3, 3
d	3, -2	-2, 0

(e) Condition 5

S \ O	*c*	*d*
c	1, 2	-2, 3
d	3, -3	0, -2

(f) Condition 6

S \ O	*c*	*d*
c	1, 1	-2, 3
d	3, -2	0, 0

(g) Condition 7

S \ O	*c*	*d*
c	1, 2	-2, 4
d	2, -3	-1, -1

(h) Condition 8

S \ O	*c*	*d*
c	1, 1	-2, 2
d	2, -2	-1, -1

† The figures in the matrices indicate monetary pay-offs in Dutch cents.

a one-half replication of a 2^4 factorial design (fractional factorial design) with the fourth-order interaction of a_s, Δb_s, a_o and Δb_o as the defining relation (Winer, 1971, p. 676). This yielded eight conditions according to the scheme of Table 2.8. Table 2.9 gives the eight resulting pay-off matrices, when the proper values for the experimental parameters are substituted in Table 2.6.

The subjects were 32 male high school students, ranging in age from 15 to 17 years. They were randomly assigned to the eight experimental conditions, four in each condition. Subjects were run in pairs. At the beginning of the experiment a written instruction was given. The subjects were told that they had to make a number of decisions in a situation in which it was impossible to communicate with each other, but in which the effect of their choice would be partly dependent on their own decision and partly on the other person's decision. Subjects were also told that after each trial one member of each pair would be informed as to which choice the other had made, whereas the other would never get this information. Actually, both subjects were told that they would be informed about each other's choice and that the other would not. The feedback the subjects received about Other's choice was prearranged and always showed that the other had made the cooperative response.

The subjects were seated one on each side of a screen, which exhibited the pay-off matrix to both subjects throughout the experiment. At each side of the screen there was a smal box carrying four switches and two lights. At the beginning of each trial, each subject had to indicate both his choice among the alternatives and a guess about the other's choice by turning on two of the switches. By this guess we could check whether the subjects had learned to expect a cooperative response from the other. The false feedback about Other's choice was indicated by the lights.

2.6 *Estimating the parameters*

The parameters of the Markov model where estimated through an iterative procedure which was basically the same as the one used by

Meeker (1971). However, estimation of the transition parameters is performed slightly differently. Further a method was developed to handle inconsistencies in the data.

As a starting assumption for the iterative procedure it is assumed that if a subject has given two successive c responses, he is in state C on the next trial, and if he has given two successive d responses, he is in state D on the next trial. Obviously, some subjects will be included wrongly in the value states, but it is assumed that this number will be relatively small and will be reduced in successive iterations. Thus,

$$P[C_n | c_{n-1} c_{n-2}] = 1 \tag{2.10}$$

and

$$P[D_n | d_{n-1} d_{n-2}] = 1 \tag{2.11}$$

With these assumptions we can estimate π and ρ directly by

$$\hat{\pi} = P[c_n | c_{n-1} c_{n-2}] \tag{2.12}$$

and

$$\hat{\rho} = P[d_n | d_{n-1} d_{n-2}]. \tag{2.13}$$

By applying some theorems from Bayesian probability theory we find that

$$P[C_{n-1} | d_{n-1} c_{n-2} c_{n-s}] = \frac{P[C_{n-1} d_{n-1} c_{n-2} c_{n-3}]}{P[d_{n-1} c_{n-2} c_{n-3}]} =$$

$$\frac{(1-\pi)*P[C_{n-1} | c_{n-2} c_{n-3}]}{P[d_{n-1} | c_{n-2} c_{n-3}]} \tag{2.14}$$

Similarly,

$$P[D_{n-1} | c_{n-1} d_{n-2} d_{n-3}] = \frac{(1-\rho)*P[D_{n-1} | d_{n-2} d_{n-3}]}{P[c_{n-1} | d_{n-2} d_{n-3}]} \tag{2.15}$$

Using (2.10), (2.11), (2.12) and (2.13), equations (2.14) and (2.15) become for the first iteration:

$$P[C_{n-1} | d_{n-1} c_{n-2} c_{n-3}] \quad = 1 \tag{2.16}$$

and

$$P[D_{n-1} \; c_{n-1} d_{n-2} d_{n-3}] \quad = 1 \tag{2.17}$$

Now, θ and ψ can be estimated (Meeker, 1971, p. 397) as

$$\hat{\theta} = \frac{P[C_n \mid d_{n-1} c_{n-2} c_{n-3}] - \pi}{1 - \rho - \pi} \tag{2.18}$$

and

$$\hat{\psi} = \frac{P[D_n \mid c_{n-1} d_{n-2} d_{n-3}] - \rho}{1 - \rho - \pi} \tag{2.19}$$

As the model decribes a Markov process with stationary transition probabilities, an equilibrium will be reached (e.g. Snell, 1965). If the process is in equilibrium then $P[C_n] = P[C_{n-1}]$ and, as is shown by Meeker,

$$P[C_n] = \frac{(1-\rho)\psi}{(1-\rho)\psi + (1-\pi)\theta} \tag{2.20}$$

The first estimates of π, ρ, θ and ψ are certainly inaccurate since they are based on the arbitrary assumptions of equations (2.10) and (2.11). With the first estimates of the parameters and $P[C_n]$ we can correct the starting assumptions from equations (2.10) and (2.11). The corrected probabilities, $P[C_n \mid c_{n-1} c_{n-2}]$ and $P[D_n \mid d_{n-1} d_{n-2}]$, will enable us to improve our estimates of the parameters. We can replace (2.10) and (2.11) with

$$P[C_n \mid c_{n-1} c_{n-2}] = \frac{P[C]*\pi^2 + P[D]*(1-\rho)*\{\psi\pi + (1-\psi)(1-\rho)\psi\}}{P[D]*\pi^2 + P[D]*(1-\rho)*\{\psi\pi + (1-\psi)(1-\rho)\}} \tag{2.21}$$

and

$$P[D_n \mid d_{n-1} d_{n-2}] = \frac{P[C]*(1-\pi)*\{\theta\rho + (1-\theta)(1-\pi)\theta\} + P[D]*\rho^2}{P[C]*(1-\pi)*\{\theta\rho + (1-\theta)(1-\pi)\} + P[D]*\rho^2} \tag{2.22}$$

Let $\alpha = P[C_n \mid c_{n-1} c_{n-2}]$ and $\beta = P[D_n \mid d_{n-1} d_{n-2}]$, then π and ρ are estimated as

$$\hat{\pi} = \frac{\beta * P[c_n \mid c_{n-1} c_{n-2}] - (1-\alpha) * P[c_n \mid d_{n-1} d_{n-2}]}{\alpha + \beta - 1} \tag{2.23}$$

and

$$\hat{\rho} = \frac{\alpha * P[d_n \mid d_{n-1} d_{n-2}] - (1-\beta) * P[d_n \mid c_{n-1} c_{n-2}]}{\alpha + \beta - 1} \tag{2.24}$$

To re-estimate θ we use

$$P[c_n | d_{n-1} c_{n-2} c_{n-3}] = P[C_{n-1} | d_{n-1} c_{n-2} c_{n-3}]*\{\theta(1-\rho)+(1-\theta)\pi\}$$

$$+P[D_{n-1} | d_{n-1} c_{n-2} c_{n-3}]*(1-\rho).$$

Note that, in general, $P[C_{n-1} | d_{n-1} c_{n-2} c_{n-3}] \neq P[C_{n-1} | c_{n-2} c_{n-3}]$; however, Meeker (1971 , p.401) used $P[C_{n-1} | c_{n-2} c_{n-3}]$. Knowing that

$$P[c_n | d_{n-1} c_{n-2} c_{n-3}]*P[d_{n-1} | c_{n-2} c_{n-3}] = P[c_n d_{n-1} | c_{n-2} c_{n-3}],$$

using (2.14) we obtain

$$\hat{\theta} = \frac{P[c_n d_{n-1} | c_{n-2} c_{n-3}] - \alpha(1-\pi)\pi - (1-\alpha)(1-\rho)\rho}{\alpha(1-\pi)(1-\rho-\pi)} \tag{2.25}$$

Similarly,

$$\hat{\psi} = \frac{P[d_n c_{n-1} | d_{n-2} d_{n-3}] - \beta(1-\rho)\rho - (1-\beta)(1-\rho)\rho}{\beta(1-\rho)(1-\rho-\pi)} \tag{2.26}$$

This procedure can be repeated until no further improvement follows A detailed description of this iterative process is given in Meeker (1971).

As can be seen from equations (2.12), (2.13), (2.18) and (2,19) the whole estimation procedure is based on the three-step dependencies, $P[c_n | c_{n-1} c_{n-2}]$ and $P[d_n | d_{n-1} d_{n-2}]$, and the four-step dependencies, $P[c_n | d_{n-1} c_{n-2} c_{n-3}]$ and $P[d_n | c_{n-1} d_{n-2} d_{n-3}]$. These quantities can be computed directly from the data. While in practice $P[c_n | d_{n-1} c_{n-2} c_{n-3}]$ and $P[d_n | c_{n-1} d_{n-2} d_{n-3}]$ can take values between 0 and 1, the interval for these probabilities within the model is smaller. From (2.28) and (2.19) one can derive that the following inequalities must hold:

$$1 - \rho < P[c_n | d_{n-1} c_{n-2} c_{n-3}] < \pi \tag{2.27}$$

and

$$1 - \pi < P[d_n | c_{n-1} d_{n-2} d_{n-3}] < \rho \tag{2.28}$$

If these requirements are not fulfilled, the change of state parameters would be smaller than zero or greater than one. Meeker did not point to these requirements, probably because his data dit not give him any problem in this respect. In analysing the data for the separate treatment combinations of our experiment we found that for two such combinations one of these requirements was not fulfilled. The following solution was used by us for this problem.

Notice that the four-step dependencies are computed as follows:

$$P[c_n|d_{n-1}c_{n-2}c_{n-3}] = \frac{n(c\text{-}c\text{-}d\text{-}c)}{n(c\text{-}c\text{-}d\text{-}c)+n(c\text{-}c\text{-}d\text{-}d)} \tag{2.29}$$

and

$$P[d_n|c_{n-1}d_{n-2}d_{n-3}] = \frac{n(d\text{-}d\text{-}c\text{-}d)}{n(d\text{-}d\text{-}c\text{-}d)+n(d\text{-}d\text{-}c\text{-}c)}, \tag{2.30}$$

where $n(.)$ is the number of the specified sequence of responses.

It is assumed that in the population of all behaviour sequences (of all subjects) the four-step dependencies do not exceed the limits as indicated in (2.27) and (2.28). Our observation are based on just a sample. Hence it may be possible that, due to sampling error, some empirical four-step dependencies become too small or too great. If the number of the four-step sequences, which constitute the ratio for a four-step dependency, is relatively small (say smaller than 10), then that empirical probability gives only a crude estimate of the population value. Specifically, the value of the dependency will change considerably, if one point is added to the numerator and/or the denominator of the ratio. For example, suppose we find $n(c\text{-}c\text{-}d\text{-}c) = 2$ and $n(c\text{-}c\text{-}d\text{-}d) = 6$.
This yields, by equation (2.29), $P[c_n|d_{n-1}c_{n-2}c_{n-3}] = 2/(2+6) = 0.25$. If we raise $n(c\text{-}c\text{-}d\text{-}c)$ by one point, the four-step dependency becomes $3/9 = 0.33$, but if we raise $n(c\text{-}c\text{-}d\text{-}d)$ by one point, the dependency becomes $2/9 = 0.22$.

The following procedure was applied by us in those cases for which the four-step dependencies exceeded the limits in (2.27) and (2.28):

(i) If $P[c_n|d_{n-1}c_{n-2}c_{n-3}]$ was greater than $\hat{\pi}$, the denominator of (2.29) was raised by one point. This was repeated until the dependency was smaller then $\hat{\pi}$.

(ii) If $P[c_n|d_{n-1}c_{n-2}c_{n-3}]$ was smaller than $(1-\hat{\rho})$, both numerator and denominator of (2.29) were raised by one point. This was repeated until the dependency was greater than $(1-\hat{p})$.

The same "correction procedure" was applied to $P[d_n | c_{n-1} d_{n-2} d_{n-3}]$. No more than one cycle was needed in any of our cases.

2.7. *Experimental results*

The experiment was run over 50 trials. Data consisted of sequences of c and d responses, one sequence for each subject. First, for each subject the proportion of c responses was computed. Some of the results are given in Table 2.10. Hartley's test for homogeneity of variance (e.g. Winer, 1971, p. 206) applied to the eight simple conditions yielded an F_{max} of 2.93, which is not significant.

Table 2.10 Means and variances for number of c responses (50 trials per subject)

Condition	Mean	Variance	Number of subjects
1	16.50	103.25	4
2	19.50	283.25	4
3	32.00	266.50	4
4	26.00	159.00	4
5	27.25	180.00	4
6	23.25	168.19	4
7	19.75	224.19	4
8	30.00	302.50	4
1+2+3+4 (a_s^+)	23.50	239.00	16
5+6+7+8 (a_s^-)	25.06	234.06	16
1+2+5+6 (Δb_s^+)	21.63	200.11	16
3+4+7+8 (Δb_s^-)	26.94	160.06	16
1+4+6+7 (Δb_o^+)	21.38	176.61	16
2+3+5+8 (Δb_o^-)	27.19	280.78	16
1+3+5+7 (a_o^+)	23.88	230.86	16
2.4.6.8 (a_o^-)	24.69	243.09	16

Each level of the four variables from Coombs' parameterization of the PDG was represented by four experimental treatment combinations.

To compare the effects of different levels of each variable, data for several quartets of treatment combinations were pooled as shown in Table 2.10.

As can be seen from Table 2.8, all pairs of quartets are matched with respect to main effects (except the one under consideration) and second-order interactions. There is no matching with respect to third-order interactions, but it is assumed that they are zero or at least very small compared to the main effects. In addition, the observable second-order interaction are completely confounded (aliased) with the interaction effect of the complementary pair of variables (Winer, 1971, p. 676).

The numbers of c responses were analysed by means of an analysis of variance which is the usual way which PDG studies have been analysed.

The results of the analysis of variance, shown in Table 2.11, were striking. Not only did we find no significant effects, but all F-ratios turned out to be smaller than 1.
Taking the reciprocals of the F-ratios and interchanging the degrees of freedom also did not yield any significant value. A possible explanation for these results may be that a systematic source of

Table 2.11 Analysis of variance table for number of c responses

Source and alias	SS	d.f.	MS	F†
$a_s(\Delta b_s \times a_o \times \Delta b_o)$	19.53	1	19.53	0.07
$\Delta b_s(a_s \times a_o \times \Delta b_o)$	225.78	1	225.78	0.80
$a_o(a_s \times \Delta b_s \times \Delta b_o)$	5.28	1	5.28	0.02
$\Delta b_o(a_s \times \Delta b_s \times a_o)$	270.28	1	270.28	0.96
$a_s \times \Delta b_s(a_o \times \Delta b_o)$	258.78	1	258.78	0.92
$a_s \times a_o(\Delta b_s \times \Delta b_o)$	42.78	1	42.78	0.15
$\Delta b_s \times a_o(a_s \times \Delta b_o)$	13.78	1	13.78	0.05
Within cells	6752.23	24	281.34	
Total	7588.45	31		

† All values are far from significant. MS (within cells) is used as the denominator of F in all cases.

variation is hidden in the within-cells sum of squares. Fourth-order interactions could be the answer, but seems very unlikely. If Meeker's theory holds true, the great within-cells variability might be caused by the different initial value states of the subjects. For the probability of a certain number of c responses is highly dependent on the value state. This is intuitively clear, if one considers that the probability of a c response is greater when a subject is in state C than when he is in state D. Moreover before equilibrium is reached, the probability of a subject being in state C is greatest if on the preceding trial he was also in state C.

Next, the whole response sequences were analysed following the model of Meeker as outlined in the preceding section. Even if no differences are found in the manifest behaviour, this does not imply that the latent structure of the behaviour in the different experimental conditions is invariant. To ensure that the subjects had an opportunity to learn what the response of the other was going to be on the next trial only responses of the last 40 trials were included in the analysis.

Table 2.12 shows the three-step and four-step dependencies which form the input for the estimation procedure. The numbers in parentheses for conditions 7 and 8 are the four-step dependencies after application of the "correction procedure" described in Section 2.6. In these instances the requirement given in (2.28) was violated.

Provisional estimates of the parameters were obtained by equations (2.12), (2.13), (2.18) and (2.19). The estimates were based on the following two assumptions: (i) that after two successive c responses the subject is in C; and (ii) that after two successive d responses the subject is in D. Using these first estimates of the parameters and $P[C]$, the error in the starting assumptions was estimated by the iteration procedure as described in Section 2.6.

The estimates of the parameters rapidly converged, the process of iteration being terminated on 0.001 change for each individual parameter. In fact, for 13 out of the 16 conditions the process converged after only two cycles, and in condition 6, 7 and 8 after three, four and five cycles respectively. The final solutions, which

Table 2.12 Observed three-step and four-step dependencies over 40 trials

Conditions	$P[c_n \mid c_{n-1}c_{n-2}]$	$P[d_n \mid d_{n-1}d_{n-2}]$	$P[c_n \mid d_{n-1}c_{n-2}c_{n-3}]$	$P[d_n \mid c_{n-1}d_{n-2}d_{n-3}]$
1	0.769	0.867	0.222	0.364
2	0.900	0.944	0.750	0.750
3	0.919	0.838	0.286	0.333
4	0.855	0.831	0.250	0.454
5	0.882	0.846	0.333	0.250
6	0.855	0.863	0.625	0.429
7	0.854	0.915	0.833	1.000 (0.875)
8	0.886	0.951	0.778	0.000 (0.333)
1+2+3+4	0.873	0.875	0.321	0.438
5+6+7+8	0.873	0.894	0.625	0.500
1+2+5+6	0.857	0.883	0.433	0.400
3+4+7+8	0.885	0.884	0.533	0.538
1+4+6+7	0.837	0.872	0.452	0.528
2+3+5+8	0.897	0.900	0.517	0.350
1+3+5+7	0.872	0.874	0.387	0.469
2+4+6+8	0.873	0.895	0.586	0.458

Table 2.13 Estimates of parameters after iteration procedure

Condition	$\hat{P}[C_n \mid c_{n-1}c_{n-2}]$	$\hat{P}[D_n \mid d_{n-1}d_{n-2}]$	$\hat{\pi}$	$\hat{\rho}$	$\hat{\theta}$	$\hat{\psi}$	$\hat{P}[\ ]$
1	1.0	1.0	0.77	0.87	0.86	0.79	0.35
2	1.0	1.0	0.90	0.95	0.18	0.23	0.41
3	1.0	1.0	0.92	0.84	0.84	0.67	0.62
4	0.99	1.0	0.86	0.83	0.88	0.54	0.43
5	1.0	1.0	0.88	0.85	0.75	0.82	0.58
6	1.0	0.98	0.86	0.88	0.31	0.60	0.62
7	0.99	0.98	0.86	0.93	0.03	0.05	0.49
8	1.0	0.98	0.89	0.97	0.13	0.74	0.63
1+2+3+4 (a_s^+)	1.0	1.0	0.88	0.88	0.74	0.58	0.44
5+6+7+8 (a_s^-)	1.0	0.99	0.87	0.90	0.32	0.51	0.56
1+2+5+6 (Δb_s^+)	1.0	1.0	0.86	0.89	0.57	0.65	0.48
3+4+7+8 (Δb_s^-)	1.0	1.0	0.89	0.89	0.46	0.45	0.50
1+4+6+7 (Δb_o^+)	0.99	1.0	0.84	0.88	0.54	0.48	0.41
2+3+5+8 (Δb_o^-)	1.0	1.0	0.90	0.90	0.47	0.69	0.58
1+3+5+7 (a_o^+)	1.0	1.0	0.88	0.88	0.65	0.54	0.45
2+4+6+8 (a_o^-)	1.0	0.99	0.88	0.90	0.37	0.57	0.55

resulted from the process of iteration, are listed in Table 2.13, and in only two cases in the whole table were the initial and final estimates discrepant by more than 0.02.

As can be seen from Table 2.13, the estimated probability that a wrong subject is included in a certain value state due to one of the starting assumption (2.10) and (2.11) is zero or almost zero. To check how well the estimated parameters fitted to the data, comparisons were made between some theoretical probabilities predicted by the estimated parameters and the corresponding proportions observed in the data. None of these proportions were used in the estimation procedure. The theoretical probabilities are the expected proportions (at equilibrium) of c, c following c, and d following d. The equations for deriving these probabilities are given in Meeker (1971).

The results of the comparisons are displayed in Table 2.14. Except for conditions 6, 7 and 8, the absolute differences between the observed and predicted probabilities of c are very small (0.03 or less). For the former conditions these discrepancies are 0.08, 0.10 and 0.06 respectively. The discrepancies between the observed and predicted conditional probabilities are larger on the average and show a greater variance than do the discrepancies for $P[c]$. In particular, the results for condition 2 are unfavourable. Our results for the conditional probabilities are somewhat disappointing when compared to the experimental results obtained by Meeker (1971, p. 398). One explanation for the poor fit of the first-order conditional probabilities might be that the observed statistics are unreliable, since they are generally based on fewer observations than the $P[c]$ statistics. The estimates of the parameters in the single conditions are based on a relatively small number of observations (four sequences of 40 trials only). In particular, the data provide only a crude estimate of the four-step dependencies.

In the Appendix a proof is given showing that $P[c]$ increases as

Table 2.14 Comparisons between predicted and observed response probabilities

	$P[c]$		$P[c_n \mid c_{n-1}]$		$P[d_n \mid d_{n-1}]$	
Condition	Predicted	Observed	Predicted	Observed	Predicted	Observed
1	0.35	0.35	0.74	0.73	0.86	0.85
2	0.40	0.40	0.85	0.65	0.90	0.78
3	0.63	0.66	0.89	0.85	0.82	0.74
4	0.47	0.48	0.80	0.77	0.82	0.82
5	0.58	0.58	0.87	0.86	0.82	0.82
6	0.58	0.50	0.83	0.71	0.77	0.70
7	0.46	0.36	0.80	0.74	0.83	0.86
8	0.57	0.63	0.88	0.84	0.84	0.72
1+2+3+4	0.46	0.47	0.83	0.77	0.86	0.81
5+6+7+8	0.53	0.52	0.84	0.79	0.82	0.78
1+2+5+6	0.47	0.46	0.83	0.75	0.84	0.79
3+4+7+8	0.50	0.53	0.84	0.81	0.84	0.80
1+4+6+7	0.42	0.42	0.78	0.74	0.84	0.81
2+3+5+8	0.56	0.57	0.88	0.81	0.81	0.77
1+3+5+7	0.47	0.49	0.82	0.81	0.85	0.83
2+4+6+8	0.52	0.50	0.85	0.75	0.83	0.76

$P[d_n \mid c_{n-1}d_{n-2}d_{n-3}]$ decreases and also as $P[c_n \mid d_{n-1}c_{n-2}c_{n-3}]$ increases. It can also be shown that changing the four-step dependencies in such a way as to yield parameters which predict a greater $P[c]$ will also cause an increase in $P[c_n \mid c_{n-1}]$ and a decrease in $P[d_n \mid d_{n-1}]$. However the change in the conditional probabilities will be smaller than in $P[c]$. This means that this procedure only makes sence, if its main purpose is to improve the estimate of $P[c]$.

Both four-step dependencies in conditions 6, 7 and 8 were changed in the stepwise manner described earlier. After these "corrections" new estimates of the parameters were computed as shown in Table 2.15. Table 2.16 shows the new predicted probabilities. Clearly the new predictions show less discrepancy with the observed statistics than did the earlier ones.

Table 2.15 Estimates of parameters after adjustment of the four-step dependencies

Condition	$P[C_n \mid c_{n-1} c_{n-2}]$	$P[D_n \mid d_{n-1} d_{n-2}]$	$\hat{\pi}$	$\hat{\rho}$	$\hat{\theta}$	$\hat{\psi}$	$P[C]$
6	1.0	0.99	0.86	0.87	0.41	0.50	0.52
7	0.99	0.99	0.86	0.92	0.13	0.15	0.38
8	1.0	0.98	0.89	0.97	0.10	0.84	0.68

Table 2.16 Comparisons between predicted and observed probabilities for revised estimates of parameters

	$P[c]$		$P[c_n \mid d_{n-1}]$		$P[d_n \mid d_{n-1}]$	
Condition	Predicted	Observed	Predicted	Observed	Predicted	Observed
6	0.51	0.50	0.81	0.71	0.80	0.70
7	0.38	0.36	0.78	0.74	0.87	0.86
8	0.62	0.63	0.88	0.84	0.81	0.72

In light of the fact that the procedure of estimating the parameters on which the predictions are based, was not derived in such a way as to maximize the fit of the model to the data (e.g. least squares, maximum likelihood, or minimum chi-squared procedures), at least for the joint conditions the predictions seem satisfactory. As a starting point in our iteration procedure we used the assumptions from (2.10) and (2.11). In most cases the assumptions turned out to be right, but one may wonder wether the solutions found for the parameters are unique. Would not we have found other estimates for the parameters, if the starting assumptions had been different?

To examine this possibility we changed the probabilities in (2.10) and (2.11) to 0.90, 0.80 and 0.70 successively. This implied that the first estimates of the four parameters had to be computed using equations (2.23), 2.24), 2.25) and (2.26) instead of (2.12), (2.13), (2.18) and (2.19). Next the iterative process was entered with these (new) first estimates for the four parameters. In all conditions and for all values of the probabilities in (2.10) and (2.11) the iterative process yielded exactly the same solutions as in the orginal calculation. This indicates that the assumptions in (2.10) and (2.11) are a good starting point for the estimation procedure. Moreover, in some cases, when the probabilities in (2.10) and (2.11) were set a 0.80 or 0.70, the first estimates for π and or ρ became even larger than 1.0.

2.8. *Discussion*

Analysis of variance on the proportion of c choices did not reveal any significant effect. However, the estimates of the four parameters do vary over the conditions. This means that different reward structures which, in the average number of c choices showed no effect, in fact cause different intrapersonal conflicts.

The first hypothesis said that π and ρ, the probabilities of consistent behaviour in the value states, are not affected by variations in the pay-off structure. In our experiment π and ρ remained fairly constant over the joint conditions, but were less obviously constant in the single conditions (see Table 2.13). As noted earlier, however, we may assume that the estimates for the four parameters are more reliable in the joint conditions than in the single conditions, where we have fewer observations. For the time being we suspend judgement about the tenability of this hypothesis until more data are available.

The results with respect to θ and ψ, the change of state parameters, agree with the relevant hypotheses rather well. From Table 2.13, it follows that θ, the probability to change from C to D after an inconsistent choice, is very sensitive to changes in a_s and a_o. Both these variables indicate how much influence the players have on the outcome of the other. The fact that they influence mainly θ, which is a measure of the effect behaviour has on state C, supports the theoretical assumption that a subject in state C is directed towards the other. There is also a slight effect of Δb_s on θ. This can be explained as follows: as the inducement to defect (Δb_s) increases, the psychological cost for subject to change his value state from C to D becomes smaller.

Wit respect to ψ, the probability of changing from D to C after inconsistent behaviour, we can say that the results are in favour of the hypotheses. First, the data show no effect (or a negligible effect) of a_s and a_o on ψ. This is in agreement with the definition of state D, the rational state: the subject is directed towards his own outcome solely, and not towards possible costs to others caused by his behaviour. Since a_s is positively present in the outcome of either choice [(a_s+b_s) for the c choice and (a_s+b_s+Δb_s) for the d choice], it is not surprising that it does not affect the rational state.

Secondly, there is a considerable increase in ψ, as Δb_s increases. This result was expected from the model, which states that in state D the subject is purely rational and therefore the larger the cost attached to a c choice for subject, the more conflicting cooperative behaviour is when the subject occupies state D. When the subject is in value state D he will not only judge his own alternatives rationally, but the whole outcome structure, including the possible outcomes for the other. The transition from D to C will come about easier (ψ will become larger) when Δb_o becomes smaller. This can be interpreted as follows: when the other has less opportunity to benefit himself (as expressed by monetary amounts), the tendency to cooperate will increase in the subject. Probably, Δb_o has the same function in the transition from D to C as Δb_s has in the transition from C to D, let alone that the effects are in opposite directions.

In the estimation procedure response sequences were pooled to get more reliable estimates of the three- and four-step dependencies. Besides, not all dependencies could be estimated from some single response sequences (e.g. in condition 2 one subject never cooperated). One may object that there is no evidence that the individual chains are homogeneous. However, if they are not, this would be indicated by the results of the analysis, since addition of heterogeneous chains does not, in general, result in a new chain.

Although a test for the significance of the results is lacking, we feel that the results clearly demonstrate the effects which we expected to appear. At least this is true with respect to θ and ψ.

The results provide new support for Meeker's theory which makes his approach to the analysis of sequential behaviour promising. The most important feature of the Markov chain model, as constructed by Meeker, is that it makes transition from the one (latent) state to the other contingent upon the subject's behaviour. Meeker constructed his theory for a very specific hypothetical situation. In our view, research on his theory has to direct itself to expanding the applicability of the theory. One could start with those cases, in which Other is not always cooperative. It would be interesting to study the effect of different probabilities with which Other chooses the cooperative act on the four parameters. It is our hypothesis that this will affect the values of π and ρ. We leave this possibility for future experiments.

References

Coombs, C.H. (1973. A reparameterization of the Prisoner's Dilemma Game. *Behavl Sci. 18*, 424-428.

Friedland, N., Arnold, S.E. & Thibaut, J. (1974). Motivational bases in mixed-motive interactions: The effects of comparison levels. *J. exp. Soc. Psychol. 10*, 188-199.

Griesinger, D.W. & Livingston, J.W., Jr. (1973) Toward a model of interpersonal motivation in experimental games. *Behavl. Sci. 18*, 173-188.

Kuhlman, D.M. & Marshello, A.F.J. (1975). Individual differences in game motivations as moderators of preprogrammed strategy effects in Prisoner's Dilemma. *J. Person. Soc. Psychol. 32*,922-931.

Luce, R.D. & Raiffa, H.(1957). *Games and Decisions*. New York: Wiley.

Meeker, B.F. (1971) Value conflict in social exchange: A Markov model. *J. Math. Psychol. 8*, 389-403.

Messick, A.M. & McClintock, C.G. (1968). Motivational bases of choice in experimental games. *J. exp. soc. Psychol. 4*, 1-25.

Pruitt, D.G. (1970) Motivational processes in the decomposed Prisoner's Dilemma Game. *J. Person. soc. Psychol. 14,* 227-238.

Rapoport, A. & Chammah, A. (1965) *Prisoner's Dilemma.* Ann Arbor: University of Michigan Press.

Scodel, A., Minas, J., Ratoosh, P. & Lipetz, M. (1965), Some decriptive aspects of two-person non-zero-sum games. *J. Conflict Resolution. 3,* 114-119.

Snell, J.L. (1965) Stochastic processes. In: R.D. Luce, R.R. Bush & E. Galanter (Eds.), *Handbook of Mathematical Psychology,* vol. III, chapter 20. New York: Wiley.

Winer, B.J. (1971) *Statistical Principles in Experimental Design,* 2nd ed. New York: McGraw-Hill.

Wyer, R.S., Jr. (1971) Effects of outcome matrix and partner's behavior in two-person games. *J. exp. soc. Psychol. 7,* 190-210.

Appendix

From (2.20) and axiom 2-2 it follows that

$$P[c] = \frac{(1-\rho)\psi\pi+(1-\pi)*\theta*(1-\rho)}{(1-\rho)*\psi+(1-\pi)*\psi} \tag{2.31}$$

Now, the partial derivatives of $P[c]$ for θ and ψ are

$$\frac{\partial P[c]}{\partial\theta} = \frac{(1-\pi)(1-\rho)\{(1-\rho)\psi+(1-\pi)\theta\}-(1-\pi)\{(1-\rho)\psi\pi+(1-\pi)(1-\rho)\theta\}}{\{(1-\rho)\psi+(1-\pi)\theta\}^2}$$

$$= \frac{(1-\pi)(1-\rho)(1-\rho-\pi)\psi}{\{(1-\rho)\psi+(1-\pi)\theta\}^2} < 0 \tag{2.32}$$

$$\frac{\partial P[c]}{\partial\psi} = \frac{(1-\rho)\pi\{(1-\rho)\psi+(1-\pi)\theta\}-(1-\rho)\{(1-\rho)\psi\pi+(1-\pi)(1-\rho)\theta\}}{\{(1-\rho)\psi+(1-\pi)\theta\}^2}$$

$$= \frac{(\pi+\rho+1)(1-\rho)(1-\pi)\theta}{\{(1-\rho)\psi+(1-\pi)\theta\}^2} < 0 \tag{2.33}$$

So $P[c]$ increases when ψ increases and decreases when θ increases. Holding π and ρ constant, ψ will increase if $P[d_n|c_{n-1}d_{n-2}d_{n-3}]$ decreases. This can be seen easily in equation (2.19).

From equation (2.18) it follows that π and ρ again being equal, θ will increase if $P[c_n|d_{n-1}c_{n-2}c_{n-3}]$ decreases. Now it is clear that an increase in $P[d_n|c_{n-1}d_{n-2}d_{n-3}]$ is accompanied by a decrease in $P[c]$.

Raising the value of $P[c_n | d_{n-1} c_{n-2} c_{n-3}]$ will cause the opposite effect in $P[c]$.

Note: The derivations in this appendix are presented with (2.18) and (2.19) as references for the equations for θ and ψ. The derivations are, however, also valid for the exact equations (2.25) and (2.26) for θ and ψ. Moreover, the data in the first two columns of Table 2.13 are an extra justification for the use of the simplified formulas.

3.1. *Introduction*

The latent Markov model developed by Meeker (1971) was designed to describe the process of decision making by an individual who must make a series of choices between reciprocating the cooperative acts of another individual, or maximizing his own payoff in a situation where the other does not retaliate (i.e. is unresponsive to the subject's behavior).

Meeker as well as Van der Sanden (1978, see also chapter 2) in their experiments used the highly restrictive condition that the other player (O) in the experimental game was always cooperative towards the subject (S), that is O always chose the cooperative alternative in a PDG-type situation (see Table 2.1).

Since the original model was succesful in describing the actual behavior in experiments under the above mentioned restricted conditions it seems worthwhile to investigate whether the model can be extended to more complex situations. In this chapter a generalization will be presented to situations where

1. both actors (S and O) in the social exchange situation have k choice alternatives ($k \geqslant 2$), and
2. O chooses each alternative with a fixed (constant) probability, non-contingent with subject's behavior, and O is no longer always cooperative.

In Meeker's theory (which applies to situations where a subject can only choose between two acts a cooperative act and a non-cooperative act) it is assumed that during the decision making process the subject is in one of two, so called, *value states*. A value state is defined as the set of values in the social exchange situation which give the same prescriptions as to what is the most appropriate behavior. Each value state has its own probability distribution over the choice alternatives open to the subject. The behavior which is "the most appropriate" in a particular value state, also has the greatest probability in that value state and is called *"consistent behavior"* for that value state.

When a subject must make a series of choices between cooperating and non-cooperating, it is assumed that he may change his value state during the process as a result of his own behavior and his opponent's behavior according to some (probabilistic) transition rules.

Value states influence behavior, but the actual behavior (of all actors involved in the interaction) also has its effect on the value states. An outcome, defined as a pair of responses by S and O, which agrees with Subject's value state, gives him (S) no reason to change his value state. Such an outcome is called a *"consistent outcome"* for that value state. In the Prisoner's Dilemma Game the outcome (*cc*),in which both players choose the cooperative alternative, is a consistent outcome for a subject in the Reciprocating State. The outcome (*dc*) in which S chooses the "rational" alternative and O the "cooperative" alternative, is a consistent outcome for S, if he occupies the rational State. For in (*dc*) S maximizes his own payoff (at O's expense indeed).

Now the following axioms about value states can be formulated:

Axiom 3-1: If a person S is engaged in interaction with another person O, he (S) occupies on any trial n a value state characterized by the values consistent with one choice alternative.

Axiom 3-2: On any trial n, S chooses one alternative. The alternative with which the value state of S on trial n is consistent, is given with the greatest probability relative to the other available alternatives. 1)

Axiom 3-3: Outcomes in the interaction between S and O are defined as pairs of responses, one response by S and one by O. Some outcomes are defined as consistent with a value state. If the outcome on any trial n is consistent with the value state which S occupies on trial n, S will continue to be

1) In other words, the conditional response distribution for each value state is unimodal.

in that value state on the next trial. Otherwise, S may change to another value state on the next trial.

3.2. *Formulating the mathematical model*

Definition 1. Let Y be the set of choice alternatives (responses) open to the Subject. On each trial exactly one of the elements of Y is observed. The set Y consists of k distinct elements:

$$Y = (y_1, y_2, \ldots, y_k) \tag{3.1}$$

Y is called the set of observable states.

Note: A set of choice alternatives is likewise defined for the Other. Although it is assumed here that O has the same number of choice alternatives (k), this is not essential for the theory. O may as well have more or less choice alternatives than S.

Definition 2. U is the set of latent states (value states). On each trial the subject is in exactly one of these states. The set U also has k distinct elements:

$$U = (u_1, u_2, \ldots, u_k) \tag{3.2}$$

The model is limited to an equal number of latent and observable states. However, if on theoretical grounds these numbers are assumed to be different, it can still be applied by combining either underlying (latent) states or observable states.

Definition 3. Q is the (k×k)-matrix of response probabilities with entries:

$$q_{ij} = \text{Prob}\,[\,y_j | u_i\,] \tag{3.3}$$

According to *Axiom* 3-2

$$\text{Prob}\,[\,y_i | u_i] > \text{Prob}\,[\,y_j | u_i] \quad \text{for } i,j = 1, \ldots, k \quad \text{and } i \neq j$$

Definition 4. The (manifest) outcome on each trial is determined by Subject's response and Other's response. Let Ω be the set of outcomes. Then Ω is the Cartesian product of Y_s and Y_o. (We assume $Y_s = Y_o$.) Each element ω_m is an outcome resulting from S choosing y_j and O choosing $y_{j'}$, $(j, j' = 1, \ldots, k)$ with

$$m = (j-1) * k + j', \qquad m = 1, \ldots, k^2 \tag{3.4}$$

Definition 5. When a subject is in state u_g on trial n and chooses alternative y_j on trial n, and the other chooses alternative y_j', the subject will change his latent state from u_g to u_h on trial n+1 with probability $t_{gh}^{(m)}$.

This yields as many transition matrices as there are outcomes (k^2). Each transition matrix, T_m, $m=1, \ldots, k^2$, has entries

$$t_{gh}^{(m)} = \text{Prob}\,[\, u_{h_{n+1}} | u_{g_n} \;\; \omega_{m_n} \,] \tag{3.5}$$

Let B be the relation "*is consistent with*" defined on $U \times \Omega$, that is $B(u_g, \omega_m)$ if and only if ω_m is consistent with u_g.
Then according to *Axiom* 3-2

$$\text{Prob}\,[\, u_{h_{n+1}} | u_{g_n} \;\; \omega_{m_n} \,] = 0 \qquad \text{if } B(u_g, \omega_m) \quad \text{and } g \neq h$$

and

$$0 \leqslant \text{Prob}\,[\, u_{h_{n+1}} | u_{g_n} \;\; \omega_{m_n} \,] \leqslant 1 \qquad \text{if not } B(u_g, \omega_m)$$

Definition 6. The latent Markov process is assumed to be irreducible and aperiodic (that is, it is possible to get from any state to any other state in a finite number of trials). Thus, for $n \to \infty$, the process will eventually reach an equilibrium. When the process is in equilibrium Prob (u_{i_n}) = Prob $(u_{i_{n+1}})$.

The vector $\underline{v}$ has elements v_i, $(i = 1, \ldots, k)$, according to

$$v_i = \text{Prob}\,[u_{i_n}] \qquad \text{for } n \to \infty \tag{3.6}$$

Since the initial or a priori state distribution is unknown, $(k-1)(k^3 + k + 1)$ latent variables have to be identified:

$(k-1)$ variables of the type $\text{Prob}\,[u_{i_n}]$

$k(k-1)$ variables of the type $\text{Prob}\,[y_{j_n} | u_{i_n}]$

$k^3(k-1)$ variables of the type $\text{Prob}\,[u_{h_{n+1}} | u_{i_n}\ \omega_{m_n}]$

These variables are summarized in the equations for the observable second-order probabilities:

$$\text{Prob}\,[y_{g_{n+1}}\ \omega_{m_n}] =$$

$$\xi_{j'_n} * \sum_{i=1}^{k} \{\text{Prob}\,[u_{i_n}] * \text{Prob}\,[y_{j_n} | u_{i_n}] *$$

$$\sum_{h=1}^{k} (\text{Prob}\,[u_{h_{n+1}} | u_{i_n}\ \omega_{m_n}] * \text{Prob}\,[y_{g_{n+1}} | u_{h_{n+1}}])\},$$

where $m_n = (j_n - 1) * k + j'_n$

and $j_n, j'_n, g_{n+1} = 1, \ldots, k.$

When the process is in equilibrium, we can write

$$\text{Prob}\,[y_{g_{n+1}}\ \omega_{m_n}] = \xi_{j'} * \sum_{i=1}^{k} \{v_i * q_{ij} * \sum_{h=1}^{k} (t_{ih}^{(m)} * q_{hg})\}$$

Clearly, this yields k^3 equations in $(k-1)(k^3+k+1)$ unknowns. For $k \geqslant 2$ this system is underdetermined. Therefore, to solve for all the unknowns of the model one has to make use of third-order or higher-order probabilities.

3.3. *Problems relating estimability*

The model presented in the last section is a special case of the class of models in which a set of observable behaviors is dependent upon a number of underlying organismic states. It is assumed that the distribution of responses (choices) changes over time due to (outcome-contingent) transitions among hypothesized motivational states.

The model is in the form of a latent Markov chain model. Markov chain models with latent states are not new (Atkinson, Bower and Crothers, 1965; Greeno, 1974; Greeno and Steiner, 1964; Lazarsfeld and Henry, 1968; Nahinsky, 1973), although specific models quite often generate specific problems with respect to the estimation procedures.

In the (general) value state model of the last section the latent states neither can be made directly observable (through a one-to-one correspondence with observable responses for example) nor can be inferred a posteriori from the data (for example, in most learning models the hypothetical states of the subject are identified retrospectively through the assumption of learning as an absorbing event and through the use of learning criteria for the responses). Estimation in our latent Markov chain model was accomplished through a similar process as the one proposed by Meeker (1971) and applied by Van der Sanden (1978) (see chapter 2).

Alternative estimation methods, frequently applied in the context of Markovian models, were also considered but they were discarded due to the special characteristics of our model. In the next sections this will be explained in more detail.

3.3.1 *Markov models with observable states*

Early mathematical models of social interaction have as the central variables not indices of "psychological states", but rather whatever happens to be easily and obviously quantifiable, such as easily identifiable acts, which can be quantified in terms of temporal relative frequency (Rapoport, 1963, p.567). This approach has been proved to be especially fruitful in the application of mathematical models to simple learning processes, where the "state" of the subject at any time is defined by the response to which a single stimulus is conditioned at the same time. That response is given on the next trial, and so the state of the subject is made observable. An application of this type of model to two-person interaction can be found in Burke (1959).

Revenstorff, Wegschneider, Fitting and Mai (1974) have described how Markov models can be applied to nonzero-sum games and what make them (that is Markov models) so attractive. Markov models require only simple assumptions about the decision making organism (small number of parameters, simple temporal dependencies and constancy of the process). Moreover, Markov models are mathematically tractable. Very frequently simple maximum likelihood estimators for the parameters are derived and, if the solution for a Markov model is known, several descriptive quantities can be derived, which are very informative of the behavior of the system represented by the Markov model (Kemeny and Snell, 1960; see also chapter 5).

A Markov chain is a Markovian system (that is transition from one state to another is dependent only upon the last state) with stationary transition probabilities (that is transition probabilities are independent of trial number). Suppes and Atkinson (1960) applied Markov chains as models of social interaction to experimental games. Starting from axioms about behavior mechanisms based on stimulus-sampling theory they applied Markov chain models to the conditioned states of the players to describe their (i.e. the players') joint game playing behavior as learning events.

Other applications of Markov chain models to experimental games, particularly the Prisoner's Dilemma Game, are known from Rapoport and Chammah (1965) and Rapoport and Dale (1966). In these models the manifest outcomes of the game were taken as the states of the Markov process.

Estimation procedures for Markov chain models of manifest behavior are rather straightforward (see for example Anderson and Goodman, 1957; Madansky, 1959; Billingsley, 1961) since all statistics such as the vector of initial state probabilities and the transition probabilities, are computed readily from the experimental data. Many learning models are of the Markovian type (Greeno, 1974). The major estimation methods for these models are the method of moments, (including the modified method of moments), maximum likelihood estimation, minimum chi-square techniques and the method of least squares. Detailed reviews of these methods and their applications to Markov chains are found, for example, in Atkinson *et al.* (1965), Bush (1963) and Restle and Greeno (1970).

Simple as observable Markov models may be, they frequently suffer from a number of drawbacks: for instance they lack theoretical foundation, that is they are pragmatic, and conclusions about the psychological processes governing subjects' behavior (such as motivations, expectations) can be derived only a posteriori.

3.3.2. *Markov models with nonobservable states*

When a model relates observable behavior to a set of hypothetical states, which are not directly observable, things are not so easy. In this section we shall demonstrate which complications arise, when the above mentioned estimation methods are to be applied to a latent Markov chain.

The latent Markov chain model of Lazarsfeld and Henry (1968, ch. 9) is representative for a class of models which apply to sociological phenomena and sociometric experiments, such as attitude change and changes in voting behavior in large populations (Coleman, 1964; Wiggins, 1955) and changes over time in the configuration of interpersonal relationships (Katz and Proctor, 1959).

An essential feature of Lazarsfeld and Henry's model is, that transitions from one (latent) state to another are not contingent upon the (observable) outcomes of the process on the last trial, that is the transition matrix of the latent Markov chain is fixed over trials. Another difference between the latent Markov chain model of Lazarsfeld and Henry and our value state model is, that in the former model the initial state distribution in known.

A solution of their latent Markov chain was set forth by Lazarsfeld and Henry, which was based on (simple) matrix operations on second order and third-order manifest probabilities. Only a small number of trials is required to estimate the model's parameters. On the other hand, since a reasonable number of observations is required, the number of sample subjects must be sizable. This is an argument, which reduces the model's usefulness for experimental research.

Lazarsfeld and Henry (1968, p. 253) demonstrated, that if a subject behaves according to a latent Markov chain, the observable behavior (responses) is not a Markov chain.

Let $P_{s,t}$ be the matrix of joint probabilities fore time s and time t, then in Lazarsfeld and Henry's model this matrix can be written as

$$P_{s,t} = Q'V_s QR^{t-s}, \qquad t > s,$$

where Q = the matrix of response probabilities for the latent states (this matrix is similar to matrix Q in *Definition* 3 of Section 3.2),

V_s = the diagonal matrix of the latent distribution at time s,

R = $Q^{-1}MQ$ (M is the transition matrix of the latent Markov chain)

A similar equation for an ordinary Markov chain would be:

$$P_{s,t} = P_s R^{t-s},$$

where P_s is the diagonal matrix of the manifest probabilities at time s.

Since in general $Q'V_sQ$ will not be a diagonal matrix, this shows that the manifest behavior is not a Markov chain.

Lazarsfeld and Henry gave the following interpretation to the entries of $Q'V_sQ$: the i,j-th entry represents the probability of giving response i and response j at time s simultaneously, if one could or had to give two responses to the same question or to the same stimulus, under the assumption of local independence.

If a theory links a number of observable states to a set of underlying states, the problem is how to identify underlying structural characteristics through the use of observed distributions. This problem is known as the identifiability problem.

The problem has been considered in relation to Markovian learning models by Greeno and Steiner (1964). According to Greeno and Steiner a state specified in a theory is identifiable in the outcome-space of an experiment, if and only if for each response sequence in this outcome-space every occurrence of the state can be identified. If all of the states specified in the theory are identifiable in the outcome-space, the theory is an identifiable theory in that space.

Greeno and Steiner showed that for theories, which are not completely identifiable, a second equivalent theory may be constructed which is a Markov process with identifiable states and stationary transition probabilities. It was demonstrated by Greeno and Steiner (1964) that a three-state Markov chain with a single absorbing state is equivalent to several formalizations of all-or-none learning theories (Bower, 1961; Estes, 1960; Restle, 1962).

In general the development of the identifiable theory is only useful for simple and very small experiments (two states, three or four trials), since the states of the equivalent theory are defined in terms of possible sequences of trial-outcomes in the experiments (Greeno and Steiner, 1964). Otherwise the model becomes unwieldy. The theory can be reduced towards a simpler form by lumping a number of states (for a definition of lumpability see for example Kemeny and Snell, 1960). A necessary and sufficient condition for lumpability is the occurence of reccurent events. In many learning theories errors are reccurent events. It appears that the existence of reccurent events depends upon the presence of one or more absorbing states in the original theory.

Our theory under discussion clearly is unidentifiable in the sense of Greeno and Steiner, and since no reccurent events can be specified, an identifiable theory in the sense of Greeno and Steiner cannot be constructed. When an identifiable theory could be constructed, which is equivalent to the original unidentifiable theory, a number of techniques are available to estimate the parameters of the resulting models. Frequently, moment estimates are easily derived (Polson, 1970; Greeno, 1974).

A very attractive method of estimation is the method of maximum likelihood, which was introduced by R.A. Fisher in 1921.
For a general discription of this method see, for example, Bush (1963) and Wani (1971). The method consists of expressing the probability of a random variable (the observations) as a function of the parameter(s) to be estimated (the likelihood function), maximizing this function and then solving for the parameter(s). Provided maximimum likelihood estimators exist, the determination of ML estimators is relatively simple if the likelihood function is twice differentiable in the entire domain of the possible values of the parameters to be estimated (the regular case). Special techniques are required if the range of the random variable involved depends upon the parameters to be estimated (irregular case; see also Wani, 1971, p. 178-182). In the case of a Markovian process it would be straightforward to write out all possible outcome sequences of an experiment in terms of the parameter(s) to be estimated, construct the likelihood function and maximize this function (by numerical methods).

For instance, for our value state model maximum likelihood estimation could be applied to all 4-tuples (remember that at least k^4 equations are needed), but even this would present enormous computational problems. Unless restrictions on the unknown parameters are imposed, every 4-tuple could be generated by k^4 sequences of latent states, thus making the equations for these 4-tuples very complicated. Besides, the number of subjects needed to yield enough data for parameter estimation would be very large (a multiple of k^4). If only a small number of subjects is available, the problem could be circumvented by running these subjects for a great number of trials, then cutting up the total outcomes sequences in 4-tuples, pooling these 4-tuples and apply maximum likelihood estimation. Even if this method could be justified by proper assumptions in the model, it still leaves the problem of yielding bizarre equations.

The same arguments as those mentioned against the appropriateness of maximum likelihood estimation apply when we consider the minimum chi-square method of parameter estimation for the 4-tuples (for a discussion of this technique for simple Markov chains see Atkinson et al., 1965).

Finally, we want to mention a study by Nahinsky (1973), who treated the identifiability problem in a somewhat different context. He discussed the problem of identifying probabilities for sequences involving k underlying states for experiments in which there are k observable states, with the conditional distributions of observable behaviors for the underlying states specified (known or hypothezised). Nahinsky investigated the conditions under which the underlying distribution can be estimated without postulating an a priori restriction upon the underlying state distributions and an a priori restriction upon the transition rules that govern the changes from one state to another. Strictly speaking 'identifiability' for Nahinsky is equivalent to 'estimability'.

Solutions, in the form of iterative matrix methods, were suggested by Nahinsky for both outcome-noncontingent and outcome-contingent models, which require that starting vectors and conditional response probabilities are known. This leaves Nahinsky's study of no use for our model.

3.4. *Estimating the parameters*

Traditional methods of estimation for Markov chains are based on frequency counts (Billingsley, 1961) for the state transitions. In the Markov model under consideration, however, these transitions cannot be observed.

The parameters of the model (the latent probabilities) are estimated by an iterative process. Since nothing in known about the latent states, the method starts with making the unobservable states observable through an (erroneous) starting assumption. This idea is adopted from Meeker (1971). With this starting assumption estimates for the parameters are computed. Next the starting assumption is adjusted and estimating reitcrated.

Let D be a $(k \times k)$ matrix of observable three-step dependencies with cell entries

$$d_{il} = \text{Prob}\,[\,y_{l_n} \mid y_{i_{n-1}}\; y_{i_{n-2}}\,] \tag{3.7}$$

Matrices R_m, $(m = 1, \ldots, k^2)$, are defined containing observable four-step dependencies with entries

$$r^{(m)}_{gh} = \text{Prob}\,[\,y_{h_n} \mid \omega_{m_{n-1}}\; y_{g_{n-2}}\; y_{g_{n-3}}\,] \tag{3.8}$$

A matrix Z is defined with elements z_{ji} such that

$$z_{ji} = \text{Prob}\,[\,u_{i_n} \mid y_{j_{n-1}}\; y_{j_{n-2}}\,] \tag{3.9}$$

The estimation process is started with choosing initial values for the elements of Z.
From (3.3), (3.7) and (3.9) follows

$$D = Z\,Q \Rightarrow Q = Z^{-1}\,D \tag{3.10}$$

If the process is started with the identity matrix as a first estimate of Z, Q (the matrix of response probabilities) is estimated as

$$Q = D$$

The estimation of the transition matrices T_m is a little bit more complicated. We can analyze (3.8) as

$$r^{(m)}_{gh} = [\sum_{1}\{\sum_{i}(\text{Prob}\,[u_{i_{n-1}} | \omega_{m_{n-1}}\, y_{g_{n-2}}\, y_{g_{n-3}}] * \text{Prob}\,[u_{1_n} | u_{i_{n-1}}\, \omega_{m_{n-1}}])\}$$

$$* \text{Prob}\,[y_{h_n} | u_{1_n}]] \,. \tag{3.11}$$

The application of the model is now restricted to situations where O chooses each action alternative with a fixed probability, non-contingent with Subject's behavior. If we let $\xi_{j'}$, be the probability of O choosing alternative $y_{j'}$, (and remembering that $m = (j-1)*k+j'$), we can write

$$\text{Prob}\,[u_{i_{n-1}} | \omega_{m_{n-1}}\, y_{g_{n-2}}\, y_{g_{n-3}}] = \frac{\text{Prob}\,[u_{i_{n-1}}\, \omega_{m_{n-1}}\, y_{g_{n-2}}\, y_{g_{n-3}}]}{\text{Prob}\,[\omega_{m_{n-1}}\, y_{g_{n-2}}\, y_{g_{n-3}}]}$$

$$= \frac{\text{Prob}\,[u_{i_{n-1}}\, \omega_{m_{n-1}} | y_{g_{n-2}}\, y_{g_{n-3}}]}{\text{Prob}\,[\omega_{m_{n-1}} | y_{g_{n-2}}\, y_{g_{n-3}}]}$$

$$= \frac{\text{Prob}\,[u_{i_{n-1}} | y_{g_{n-2}} y_{g_{n-3}}] * \text{Prob}\,[\omega_{m_{n-1}} | u_{i_{n-1}} y_{g_{n-2}} y_{h_{n-3}}]}{\text{Prob}\,[\omega_{m_{n-1}} | y_{g_{n-2}} y_{g_{n-3}}]}$$

$$= \frac{z_{gi} * \xi_{j'} * q_{ij}}{\xi_{j'} * d_{gj}} = \frac{z_{gi} * q_{ij}}{d_{gj}} \quad {}^{1)} \tag{3.12}$$

Let E_j be a diagonal matrix with the j-th column of Q on the diagonal, and let H_j be a diagonal matrix with the j-th column of D on the

1) Note that:

$$\text{Prob}\,[\omega_{m_{n-1}} | u_{i_{n-1}} y_{g_{n-2}} y_{g_{n-3}}] = \xi_{j'} * \text{Prob}\,[y_{j_{n-1}} | u_{i_{n-1}}]$$

and $\text{Prob}\,[\omega_{m_{n-1}} | y_{g_{n-2}} y_{g_{n-3}}] = \xi_{j'} * \text{Prob}\,[y_{j_{n-1}} | y_{g_{n-2}} y_{g_{n-3}}]$

diagonal. Then from (3.3), (3.5), (3.7), (3.9), (3.11) and (3.12) follows

$$R_m = H_j^{-1} Z E_j T_m Q \tag{3.13}$$

From (3.13) T_m can be estimated as

$$T_m = E_j^{-1} Z^{-1} H_j R_m Q^{-1} \tag{3.14}$$

Now a matrix W is formed as

$$W = \sum_{j=1}^{k} E_j \sum_{j'=1}^{k} \xi_{j'} T_m, \tag{3.15}$$

with $m = (j-1) * k + j'$.

W is the overall transition matrix, or the expectation of the transition matrix when the process is in equilibrium.

Under the equilibrium assumption, $\underline{v}$ can be estimated as

$$\underline{v}'W = \underline{v}' , \tag{3.16}$$

and $\underline{v}$ is estimable under the restriction $\sum_{i=1}^{k} v_i = 1$.

At this stage of the estimation process estimates for all parameters are given.

The estimation of Z can now be improved. Let

$$W_i = E_i \sum_{j=1}^{k} \xi_j T_l \tag{3.17}$$

with $l = (j-1) * k + i$,

that is: W_i is the expectation of the transition matrix when Subject

has made response y_i.
Then the i-th row of Z is computed as

$$\underline{v}'W_iW_i\,[\underline{v}'W_i\underline{e}_i]^{-1}, \tag{3.18}$$

where $\underline{e}_i$ is the diagonal of E_i written as a column vector.

Given the new values for Z we can return to (3.10) for better estimates of Q, then of the T_m's and of $\underline{v}$, and continue this procedure, iteratively, until no further improvement occurs. Although no proof that the process will converge, can be given, when testing the algorithm with artificial data it always converged to the right solution within a small number of iterations (usually less than 5 cycles).

3.5. *Experiment*

An experiment utilizing the PDG paradigm was run to test the adequacy of the model to describe data from a situation to which the theory would apply and to compare the effects of three different levels of 'cooperativeness' of the other player on the parameters of the model.

By fixing the probability of the Other choosing *c*, the theory can be applied to the Subject, when the game is iterated over a number of trials.

It is assumed that on each trial the Subject is in one of two value states: u_1 is the Reciprocating State and u_2 is the Rational State. In the Reciprocating State the Subject is motivated by norms as solidarity, cooperativeness, reciprocity, etc. In the Rational State motivations like individual rationality, self interest, and competitiveness guide Subject's behavior.

The response parameters are

$$p = \text{Prob}\,[c_n | u_{1_n}], \qquad p > .5 \tag{3,19}$$

$$r = \text{Prob}\,[d_n | u_{2_n}], \qquad r > .5 \tag{3.20}$$

I.e. p is the probability of S (subject) choosing c on trial n when he is in the Reciprocating State on trial n and r is the probability of S choosing d on trial n when he is in the Rational State on trial n (see Table 3.1).

Table 3.1 Matrix, Q, of response probabilities

	c	d
u_1	p	$1-p$
u_2	$1-r$	r

Transition parameters are represented as a and b:

$$a = \text{Prob}\,[u_{2_{n+1}} | u_{1_n}], \tag{3.21}$$

$$b = \text{Prob}\,[u_{1_{n+1}} | u_{2_n}]. \tag{3.22}$$

I.e. a is the probability of S being in the Rational State on trial n+1 after having been in the Reciprocating State on trial n, and b is the probability of S being in the Reciprocating State on trial n+1 after having been in the Rational State on trial n.

The index parameters a_i and b_i (i = 1, ..., 4) are outcome-specific transition parameters according to

$$a_i = \text{Prob}\ [\, u_{2_{n+1}} \,|\, u_{1_n}\, \omega_{i_n} \,] \,, \tag{3.23}$$

$$b_i = \text{Prob}\ [\, u_{1_{n+1}} \,|\, u_{2_n}\, \omega_{i_n} \,] \,, \tag{3.24}$$

where the outcomes ω_i $(i=1, \ldots, 4)$ are defined as

$$\omega_1 = (cc) \quad ; \quad \omega_2 = (cd) \; ;$$

$$\omega_3 = (dc) \quad ; \quad \omega_4 = (dd) \; . \tag{3.25}$$

(In this notation Subject's choice is written first!)

For each outcome of the game a matrix of transition probabilities between the value states is determined. These matrices are given in Table 3.2.

Table 3.2 Summary of outcome-specific transition matrices

		Other	
		c	d
Subject	c	ω_1 $\begin{array}{c\|cc} & u_1 & u_2 \\ \hline u_1 & 1-a_1 & a_1 \\ u_2 & b_1 & 1-b_1 \end{array}$	ω_2 $\begin{array}{c\|cc} & u_1 & u_2 \\ \hline u_1 & 1-a_2 & a_2 \\ u_2 & b_2 & 1-b_2 \end{array}$
	d	ω_2 $\begin{array}{c\|cc} & u_1 & u_2 \\ \hline u_1 & 1-a_3 & a_3 \\ u_2 & b_3 & 1-b_3 \end{array}$	ω_4 $\begin{array}{c\|cc} & u_1 & u_2 \\ \hline u_1 & 1-a_4 & a_4 \\ u_2 & b_4 & 1-b_4 \end{array}$

Table 3.3 Matrix of expected transition probabilities for the Prisoner's Dilemma when O chooses c with fixed probability ξ.

	$\langle u_1, c \rangle$	$\langle u_1, d \rangle$	$\langle u_2, c \rangle$	$\langle u_2, d \rangle$
$\langle u_1, c \rangle$	$\xi(1-a_1)p+$ $(1-\xi)(1-a_2)p$	$\xi(1-a_1)(1-p)+$ $(1-\xi)(1-a_2)(1-p)$	$\xi a_1(1-r)+$ $(1-\xi)a_2(1-r)$	$\xi a_1 r+$ $(1-\xi)a_2 r$
$\langle u_1, d \rangle$	$\xi(1-a_3)p+$ $(1-\xi)(1-a_4)p$	$\xi(1-a_3)(1-p)+$ $(1-\xi)(1-a_4)(1-p)$	$\xi a_3(1-r)+$ $(1-\xi)a_4(1-r)$	$\xi a_3 r+$ $(1-\xi)a_4 r$
$\langle u_2, c \rangle$	$\xi b_1 p+$ $(1-\xi)b_2 p$	$\xi b_1(1-p)+$ $(1-\xi)b_2(1-p)$	$\xi(1-b_1)(1-r)+$ $(1-\xi)(1-b_2)(1-r)$	$\xi(1-b_1)r+$ $(1-\xi)(1-b_2)r$
$\langle u_2, d \rangle$	$\xi b_3 p+$ $(1-\xi)b_4 p$	$\xi b_3(1-p)+$ $(1-\xi)b_4(1-p)$	$\xi(1-b_3)(1-r)+$ $(1-\xi)(1-b_4)(1-r)$	$\xi(1-b_3)r+$ $(1-\xi)(1-b_4)r$

Table 3.3 shows the transition matrix for the combinations of latent states and manifest choices, when O chooses c with fixed probability ξ.

There were three experimental conditions. In condition [90] the Other chose the cooperative alternative (c) with a probability of .90; in condition [80] this probability was .80 and in condition [60] this probability was .60.

3.5.1 *Subjects*

36 undergraduate students in Psychology participated in the experiment. They were assigned randomly to the experimental conditions, 12 in each.

3.5.2 *Procedure*

In each experimental session 6 subjects were run at a time. They sat in a room separated from each other by screens. In front of each subject was a monitor and a panel with pushbuttons. Also in front of each subject was a sheet of paper displaying the payoff matrix of Table 3.4. Subjects were told that the numbers represented monetary amounts in Dutch cents.

Table 3.4 **Payoff matrix for the experiment; the cell entries indicate amounts in Dutch cents.**

		Other: c	Other: d
Self	c	3, 3	-1, 5
	d	5, -1	1, 1

After the subjects were seated they were told that they were participating in an experiment on decision making and that they must make a series of choices, the outcome of each choice depending on their own choice as well as on the choice of one of the other participating subjects. The six subjects were then divided up into three dyads, pairs of players in a game, but they were told on each trial these pairs would be different, the assingment of the pairs being performed randomly by the computer, that also recorded subjects' behavior and controled the monitors.

The switching of partner on each trial provided for the non-contingency of other's behavior on subject's behavior.

Subjects were told that within each block of ten trials each subject was paired with each subject exactly twice. For the rest, the trials were finished off in one run. The subjects did not know the number of trials (the number was in fact 120).

Each new trial was announced through the message "Beginning of a new round. Please make your choices" appearing on the monitor. Then subjects must make their choice of an alternative in the game of Table 3.4 within 7 seconds and also predict which choice the other was going to make on that trial. Subjects were required to state their predictions about Other's choice to enable the experimenter to find out whether the subjects learned, after a while, to expect Other's cooperative choice with the correct probability.

When all subjects had indicated their choice and prediction (by pushing the proper button on the panel), feedback about the outcome on that trial was given through a message on the monitor which displayed: 1) Subject's choice, 2) Other's choice, 3) Subject's payoff, and 4) Other's payoff. This feedback was displayed for 7 seconds. After that the next trial was announced through a message on the monitor. Each subject's earnings were recorded and added, and at the end of the experimental session subjects were payed according to the total amount won.

The subjects were actually playing against a preprogrammed computer that "chose" each alternative of the PDG with a fixed probability.

3.5.3 Results

The subjects played for 120 trials. The first 20 trials were regarded as learning trials during which the subject could learn to expect the other's cooperative choice (c) with the correct probability. The patterns of the subject's predictions showed that this number was sufficient. Two subjects had to be dropped from the analysis, one subject in [80] and one in [60], since they could not be regarded as completely "naive".

Table 3.5 shows the mean and variance of the number of c responses over the last 100 trials.

Table 3.5 Mean and Variance of number of c responses over 100 trials.

Condition	Mean	Variance	Number of subjects
[90]	54.42	642.58	12
[80]	42.73	979.47	11
[60]	43.27	1101.65	11

It is not surprising that the mean number of c responses is greatest in [90] , but these gross results do not show differences between [80] and [60] . Each subject's choices on the last 100 trials were used to estimate the parameters of the latent Markov model. The protocols of subjects in the same experimental condition were pooled. [1]

The results of the estimation process for each experimental condition are given in Table 3.6.

1) The pooling of data from different subjects can be criticized, as it implies the unwarranted assumption that the same parameters hold for different subjects. On the other hand, analysis of individual data would have suffered seriously from unreliability of data on higher order dependencies. It was therefore decided to pool data, and hoped that individual differences would not seriously affect the results. The outcome appeared favorable.

Table 3.6 Parameter Estimates from Iterative Procedures.

Condition	p	r	a_1	b_1	a_2	b_2	a_3	b_3	a_4	b_4	v_1
[90]	.81	.83	.00	.72	.38	.78	.34	.00	.80	.16	.57
[80]	.73	.85	.00	.78	.00	.83	.54	.00	.29	.00	.47
[60]	.84	.87	.00	.62	.08	.42	.71	.00	.76	.11	.43

3.5.4. Discussion

A value state is a non-observable state reflecting a person's motivational orientation towards an interaction with another person. Value states influence behavior and a person may change his value state after an inconsistent outcome. The mathematical model based on the theory was applied to data from an experiment in which the Prisoner's Dilemma Game paradigm was used.

The response parameters p and r indicate the influence the value state has on behavior. In the experiments by Meeker (1971), O acted always cooperatively. Meeker found that r changed with changes in the cost of cooperative behavior and with the experimental instructions. The value of r was smallest when a group orientation was induced through the instruction. If no group orientation was induced, r was greatest when there was a (monetary) cost attached to the cooperative act. In all these cases p remained about the same. These results made Meeker conclude (1971, p. 399), that in the Rational State Subject is more oriented to himself and that in the Reciprocating State Subject is primarily oriented to the Other. Meeker then hypothesized that p will change with changes in O's behavior. In the experiment reported here p did not vary substantially although O's behavior was different in the three experimental conditions. It should be investigated whether p changes, if the behavior of a responsive Other changes. O's changing probability of cooperative bahvior had no effect upon r. This supports the interpretation of the Rational State as self-oriented and unaffected by (changes in) O's behavior.

The parameters a and b are measures for the effect of outcomes of the interaction on the value states. The solutions for a_1 and b_3

confirm our assumption formulated in *Axiom* 3.3 that subjects do not change their value state after a consistent outcome.

In effect, (cc) is a satisfying, equilibrium outcome when S is inclined to be cooperative and O generally responds likewise, and so there is no motivation for S to change his value state. Likewise, with (dd) coming up, the Subject being in an egotistic state, and the Other generally being cooperative, there is no motivation for S to change his value state.

The Subject faces, however, an inner conflict or problem when his behavior is inconsistent with his value state. We reason that this conflict is stronger when the behavior of Other and the actual outcome provide less justification for (occasionally) being untrue to oneself.

Thus, if (cc) comes up while S is in the Rational, egotistic state, and the Other is highly cooperative, the outcome would motivate the Subject to change his value orientation. If, however, the Other is only moderately cooperative (i.e. in [60], the (cc) outcome may be looked upon more easily as a chance effect and would less motivate a change of value orientation. This is expressed in the lower values for b_1 in [60] as compared to [80] and [90].

The probability of changing from the Reciprocating State to the Rational State after unilateral defection by the other (indicated by a_2) is zero and almost zero in [80] and [60]. This could be explained as follows: when the probability of cooperative behavior by the Other is considerably smaller than 1, the Subject takes a risk by choosing c, that is the risk of getting the least preferred payoff. It seems that a Subject in the Reciprocating State takes into account a 'disappointment', when the risk is considerable and consequently outcome (cd) is not (very) conflicting for him.

The situation is different for a Subject in the Rational State when confronted with (cd). For a Subject in the Rational State the probability of (cd) is relatively small in [90] and [80]. It may look to him that he missed the point completely, and he should change his value orientation (expressed in b_2). In [60], (cd) though inconsistent with an S' egotistic inclination, is nevertheless consistent with his interpretation of the game, and would therefor less motivate a change of value orientation (b_2 = .42 in [60]).

For a Subject in the Reciprocating State the outcome (*dc*) means having made an inconsistent response while the Other has made a cooperative response. In all three conditions this poses the Subject in serious conflict judging from the values of a_3 which are all clearly different from zero. a_3 is lowest in [90] and highest in [60] suggesting that it is easier to change from the Reciprocating State when the expectations of Other's cooperativeness are relatively low.

It is not surprising to find that a_4 is greater than b_4, however it is not clear why a_4 in [80] is so much smaller than a_4 in [90] and [60]. For the moment no explanation can be given for this result.

Table 3.7 Some Observed and Predicted Conditional Response Probabilities

Condition	Prob [c | (cc)]		Prob [c | (cd)]	
	Observed	Predicted	Observed	Predicted
[90]	.79	.78	.56	.58
[80]	.68	.71	.70	.72
[60]	.81	.79	.67	.72

Condition	Prob [c | (dc)]		Prob [c | (dd)]	
	Observed	Predicted	Observed	Predicted
[90]	.27	.27	.31	.28
[80]	.23	.21	.24	.24
[60]	.15	.15	.24	.22

The fit of the parameter estimates on the data was investigated by comparing some observed response probabilities, which were not used in the estimation procedure, with their predictions from the estimated parameters. As can be seen from Table 3.7 the predictions fitted the data pretty well. The average discrepancy with observed and predicted probabilities was less than .02 with the maximum discrepancy being .05.

The experiment described in this chapter is a logical extension of the experiments by Meeker (1971) and Van der Sanden (1978; chapter 2): the introduction in the social exchange situation of an "other", who is not always helpful or cooperative.

The next step would be to replace the unresponsive Other by a responsive one, changing the situation into a real interaction situation.

However, a model which captures the full complexity of the dynamics in a conflict between competition and cooperation may require some major modification of the present model.

For example, in experiments on PDG with two real players it is sometimes found that after a while players end up in some equilibrium state (Rapoport and Chammah, 1965). A Markov chain type model that can accomodate such an empirical finding must have absorbing states, the number of them being equal to the number of different alternative equilibrium states. This brings up the fundamental question of how possible equilibrium states can be distinguished beforehand.

Since some of these "equilibria" cannot be derived from the principle of rational behavior, classical game theory will not be of great help. Recently (1971) Nigel Howard has developed a theory of metagames based on the assumption that people choose conditional strategies (metastrategies) when playing experimental games. The theory enables one to derive from each game those outcomes which can be stable. The theory is predictive rather then normative and seems to be a promising development toward a better understanding of behavior in conflict situations with uncertain outcomes contingent upon Other's behavior.

Anderson, T.W. and Goodman, L.A. Statistical Inference About Markov Chains. *The Annals of Mathematical Statistics*, 1957, *28*, 89-110.

Atkinson, R.C., Bower, G.H. and Crothers, E.J. *An introduction to mathematical learning theory*. New York: John Wiley & Sons, Inc. 1965.

Billingsley, P. Statistical Methods in Markov Chains. *Ann. Math. Stat.*, 1961, *32*, 12-40.

Bower, G.H. Application of a model to paired associates learning. *Psychometrika*, 1961, *26*, 255-180.

Burke, C.J. Applications of a linear model to two-person interactions. In: R.R. Bush & W.K. Estes (Eds.), *Studies in Mathematical learning theory*. Stanford: Stanford Univ. Press, 1959 (pp. 180-203).

Bush, R.R. Estimation and Evaluation. In: R.D. Luce, R.R. Bush & E. Galanter (Eds.), *Handbook of mathematical psychology*. New York: John Wiley & Sons, 1963, Vol. I (pp. 429-469).

Coleman, J.S. *Models of change and response uncertainty*. Englewood Cliffs, N.J.: Prentice Hall, Inc., 1964.

Estes, W.K. Learning theory and the new mental chemistry. *Psychological Review*, 1960, *67*, 207-223.

Fisher, R.A. On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London*, A, ccxxii, 309-368.

Greeno, J.G. and Steiner, Th.E. Markovian processes with identifiable states: general considerations and application to all-or-none learning. *Psychometrika*, 1964, *29*, 309-333.

Greeno, J.G. Representation of learning as discrete transition in a finite state space. In: D.H. Krantz, R.C. Atkinson, R.D. Luce and P. Suppes (Eds.), *Contemporary Developments in Mathematical Psychology* (Vol. I). San Francisco: Freeman and Co., 1974.

Howard, N. *Paradoxes of Rationality: Theory of Metagames and political behavior*. MIT Press, 1971.

Katz, L. and Proctor, C. The concept of configuration of interpersonal relations in a group as a time dependent stochastic process. *Psychometrika,* 1959, *24*, 317-327.

Kemeny, J.G. and Snell, J.L. *Finite Markov Chains*. Princeton: Van Nostrand, 1960.

Lazarsfeld, P.F. and Henry, N.W. *Latent Structure Analysis*. Boston: Houghton Mifflin Company, 1968.

Madansky, A. Least squares estimation in finite Markov processes. *Psychometrika,* 1959, *24*, 137-144.

Meeker, B.F. Value conflict in social exchange: a Markov model. *Journal of Mathematical Psychology*. 1971, *8*, 389-403.

Nahinsky, L.D. Identifiablity of a class of theories with states not directly observed over trials. *Journal of Mathematical Psychology,* 1973, *10*, 296-325.

Rapoport, A. Mathematical models of social interaction. In: R.D. Luce, R.A. Bush & E. Galanter (Eds.), *Handbook of Mathematical Psychology*. New York: John Wiley & Sons, 1963, Vol. II (pp. 493-579).

Rapoport, A and Chammah, A.M. *Prisoner's Dilemma*. Ann Arbor: University of Michigan Press, 1965.

Rapoport, A and Dale, Ph. Models for Prisoner's Dilemma. *Journal of Mathematical Psychology,* 1966, *3*, 269-286.

Restle, F. The selection of strategies in cue learning. *Psychological Review,* 1962, *69*, 329-343.

Restle, F. and Greeno, J.G. *Introduction to Mathematical Psychology*. Reading Mass.: Addison-Wesley, 1970.

Revenstorff, D., Wegschneider, R., Fitting, U. and Mai, N. Markoff-Modelle für das Verhalten in experimentellen Nichtnullsummen-Spielen. In: W.F. Kempf (Ed.), *Probabilistische Modelle in der Sozialpsychologie*. Bern: Verlag Hans Huber, 1974. (in German).

Suppes, P. and Atkinson, R.C. *Markov learning models for multiperson interactions*. Stanford: Stanford Univ. Press, 1960.

Van der Sanden, A.L. Latent Markov chain analysis of a value conflict in Prisoner's Dilemma. *British Journal of Mathematical and Statistical Psychology,* 1978, *31*, 126-143.

Wani, J.K. *Probability and Statistical Inference*. New York: Appleton-Century-Crofts, 1971.

Wiggins, L.M. Mathematical models for the interpretation of attitude and behavior change. Unpublished Ph.D. - dissertation, Columbia University, 1955.

Chapter 4. Stability in the prisoner's dilemma an experimental test with general metagames

4.1. *Rational decision and game-playing behavior*

The theory of choice and decision making as a branche of pure mathematics is a theory based on axioms of consistent or rational behavior. A different word of rational behavior is optimizing behavior. A "rational decision" can be defined as follows

> "Let $D = d_1, \ldots., d_k$ be a set of alternative decisions one of wich a subject can choose and let $R = r_1, \ldots., r_k$ be a set of outcomes such that $r_j \in R$ is the forseeable result from decision $d_j \in D$. Then $d_i \in D$ is a rational decision if there is no $r_j \in R$ ($j = 1, \ldots., k$) which is preferred by the subject to the result r_i of the decision d_i. r_j is called a rational outcome."

Classical game theory (see for example Von Neumann and Morgenstern, 1944, Luce and Raiffa, 1957) is a normative theory with this Principle of Rationality as the central idea. However, particularly in the case of nonzero-sum games this principle of rationality not always results in an optimal outcome for all participants (players) in the game (chapter 1).
This problem is known as the Paradox of Rationality.

The untenability of the principle of rationality was already demonstrated in chapter 1. With respect to the Prisoner's Dilemma Game, for which frequently great discrepancies are found between the results of normative game theory and actual behavior of the players, a number of researchers has tried to solve this paradox by referring to "extra-theoretical" concepts as trust, solidarity, cooperation etcetera (Morehous, 1977, Messick and Thorngate, 1967; Messick and McClintock, 1968; Griesinger and Livingston, 1973; Friedland, Arnold and Thibaut, 1974). This indicates that we must at least reconsider the usefulness of experimental games and game theory for the study of human interaction in conflict situations. Classical, normative game theory only applies to games where the interests of the players are diametrically opposed. But theories about real human conflicts have to start at the very point where theories of zero-sum games stop. As such, nonzero-sum games are much more interesting to the behavioral scientist. As early as 1960 Schelling (1960) realized the limitations of the zero-sum game as a paradigm for human conflicts.

Rapoport (1966) has suggested two kinds of rationality: individual rationality and collective rationality. Individual rationality results in the non-optimal outcome for all players in a non - cooperative game and collective rationality prescribes the choice of the alternative which leads to the optimal outcome for all players. According to Krivohlavý (1974) the Rationality Paradox can be solved if the transition from individual rationality to collective rationality can be solved (p.596). A theory about conflict situations such as experimental games in which these meta-rational concepts are formalized and which is descriptive rather than normative is the Theory of Metagames developed by Howard (1966a, 1971).

4.2. *Meta-Games and Meta-Rationality*

The theory of metagames is based on an extension of the theoretical framework in which "rationality" is defined. As such it is a logical step in the development of theories on choice behavior from simple choice paradigms to the more complicated nonzero-sum games. This was clearly indicated by Rapoport (1967) in his popular article in Scientific American.

Howard's metagame concept which bears the "solution" to non - cooperative games such as Prisoner's Dilemma and the Game of Chicken (for a description of both games, see chapter 1) is inspired by Von Neumann and Morgenstern's idea of majorant and minorant games for zero-sum games.

Let Γ be a two-person zero-sum game. The players are labeled S_1 and S_2. The problem in analyzing a game Γ is that each player in choosing his strategy does not know what the other player is going to choose. Let Γ_1 be a game which is identical with Γ except that S_1 has to make his choice of a strategy before S_2, and that S_2 chooses his strategy in full knowledge of the choice by S_1. Clearly in this game S_1 is at a disadvantage as compared to Γ.

Let Γ_2 be a game which is identical whith Γ except that now S_2 has to make his choice of a strategy before S_1, and that S_1 chooses his strategy in full knowledge of the choice by S_2. Now, in this game, S_1 is at an advantage as compared to Γ.

Γ_1 is called the *minorant* game of Γ and Γ_2 is called the *majorant* game of Γ (Von Neumann and Morgenstern, 1944).

What do Von Neumann and Morgenstern say about how these games, Γ_1 and Γ_2, develop? "It ought to be evident by common sense - ... - that for Γ_1, Γ_2 the 'best way of playing' - i.e. the concept of rational behavior - has a clear meaning".

Let A_1 and A_2 be the sets of strategies for S_1 and S_2 resp. By definition a single play of the game consists of each player S_i ($i=1$, 2) selecting a strategy a_i from his strategy set A_i. The result is an outcome $a = (a_1 , a_2)$. A specific element from the strategy set A_i is indicated as $\bar{a}_i$, a specific outcome is indicated as $\bar{a} = (\bar{a}_1 , \bar{a}_2)$.

V_1 and V_2 are functions which attach a value to each pair of strategies (a_1 , a_2) for S_1 and S_2 resp.

In the minorant game S_1 chooses his strategy first, say $\bar{a}_1$. Now S_2 makes his choice in full knowledge of the value of a_1. It is S_2's desire to minimize $V_1(\bar{a}_1 , a_2)$ (remember that Γ is a zero-sum game). Thus, when S_1 chooses a strategy $\bar{a}_1$, S_1 can forsee with certainy what the value of $V_1(\bar{a}_1 , a_2)$ is going to be, But let us see what Von Neumann and Morgenstern have to say further.

The value of $V_1(\bar{a}_1 , a_2)$ which S_1 can forsee is Min a_2 $[V_1(\bar{a}_1, a_2)]$ Since S_1 wishes to maximize $V_1(a_1, a_2)$ and since (in Γ_1) $V_1(a_1 , a_2)$ is a function of a_1 alone (a_2 is by the principle of rationality fully determined), S_1 will choose a_1 so as to maximize Min $a_2[V_1(a_1 , a_2)]$.

This brings the value of $V_1(a_1 , a_2)$ to

$$\text{Max}_{a_1} \quad \text{Min}_{a_2} \; [V_1(a_1 , a_2)] \; .$$

How do they proceed with the majorant game? We shall omit a detailed derivation here, since Γ_2 differs from Γ_1 in that the roles of S_1 and S_2 are reversed, albeit that S_1 still wants to maximize $V_1(a_1 , a_2)$

and that S_2 still wants to minimize $V_1(a_1, a_2)$. If both players S_1 and S_2 play the majorant game well (i.e. according to the rationality principle), than the value of $V_1(a_1, a_2)$ will equal

$$\text{Min}_{a_2} \ \text{Max}_{a_1} \ [V_1(a_1, a_2)].$$

Unless $\text{Max}_{a_1} \ \text{Min}_{a_2} \ [V_1(a_1, a_2)] = \text{Min}_{a_1} \ \text{Max}_{a_2} \ [V_1(a_1, a_2)]$ (i.e. when the game has a saddle point), it is impossible for both players to be (objectively) rational in Γ. This is called by Howard (1971, p. 10) "the first breakdown of rationality". At this point Von Neumann and Morgenstern stopped their argument, but Howard elaborated their minorant and majorant games and called them *metagames*. A meta-game is the game that would exist if one of the players chose his strategy after the other(s) in knowledge of his (their) choice(s). The metagame idea was also applied to other games than zero-sum games.

Consider the Prisoner's Dilemma Game (Table 4.1). There is only one rational choice for each player: *d* ("defect"). It is the rational decision, whichever strategy the opponent chooses or is expected to choose. If the opponent is going to defect, rationally one should defect. If the opponent is going to choose cooperatively, again one should defect. The rational choice in a PDG always pays better than the irrational choice. However, if both players make the irrational choice, they will be better off then if they both choose rationally. Howard (1971, p. 48) calls this the "second breakdown of rationality".

Table 4.1 Prisoner's Dilemma Game (PDG). Ordinal preferences for each cell are shown by the number pairs. The first number in each cell shows S_1's preferences (higher numbers being more preferred) and the second number S_2's preferences.

		Player S_2 *c*	Player S_2 *d*
Player S_1	*c*	3,3	1,4
	d	4,1	2,2

The minimax strategy fails in the PDG, because the "rational" choice does not yield a jointly optimal outcome. Also, a mixed strategy is not appropriate, since each player has a dominant strategy.
If G is a game in normal form, and if k is a player in G, the k-metagame of G is the normal-form game that would exist if player k chose his strategy in G in knowledge of the other players' strategies (Howard, 1974).
The 2-metagame from the PDG is given in Table 4.2. Here player S_2 has now 4 conditional strategies, called metastrategies.

Table 4.2. 2-Metagame from Prisoner's Dilemma.
Rational outcomes for S_1 have a dot on the left, rational outcomes for S_2 have a dot on the right. The outcome (c,c), which is yielded by $(c, c/d)$ is metarational for S_1 from the 2-metagame. The equilibrium, which is a rational outcome for each player, is underlined.

S_1 \ S_2	c/c	d/d	c/d	d/c
c	3,3	1,4.	.3,3	1,4.
d	.4,1	<u>.2,2.</u>	2,2.	.4,1

These 4 metastrategies are:
1. c/c : play alternative c regardless of S_1's choice,
2. d/d : play alternative d regardless of S_1's choice
3. c/d : play alternative c or d, resp., if you think S_1 will play c or d, resp.,
4. d/c : play alternative d or c, resp., if you think S_1 will play d or c, resp.,

In the 2-metagame of Table 4.2 the outcome $(d, d/d)$ is the only equilibrium. It yields the basic outcome (d,d). S_2's rational choice is d/d and S_1's rational choice is still d. If S_1 would know that S_2 plays c/d, his best strategy is to choose c, but in that case S_2 would switch to d/d immediately. We see that $(c, c/d)$ is rational for S_1 in the 2-metagame. Hence it is said, that the basic outcome in PDG corresponding to (*c*, *c/d*), namely (*c*,*c*), is metarational from the 2-metagame for S_1. However, in the 2-metagame (*d*, *d/d*) is the only equilibrium and the dilemma still remains.

From the 2-metagame the 1,2-metagame is constructed. The 1,2-metagame is found from the 2-metagame as the 2-metagame was formed from the basic game. Thus, player 1's strategies in the 2-metagame are replaced in the 1,2-metagame by $2^4 = 16$ metastrategies. S_1 has now 2^4 metastrategies as an answer to S_2's 4 metastrategies (Table 4.3).

Table 4.3 The 1,2-metagame from Prisoner's Dilemma. The strategy "*w/x/y/z*" indicates the policy "*w* against *c/c*, *x* against *d/d*, *y* against *c/d*, *z* against *d/c*'. Rational outcomes for S_1 have a dot on the left, rational outcomes for S_2 have a dot on the right. Equilibria are underlined.

		S_1			
		c/c	*d/d*	*c/d*	*d/c*
	c/c/c/c	3,3	1,4.	.3,3	1,4.
	c/c/c/d	3,3	1,4.	.3,3	.4,1
	c/c/d/c	3,3	1,4.	2,2	1,4.
	c/d/c/c	3,3	.2,2	.3,3	1,4.
	d/c/c/c	.4,1	1,4.	.3,3	1,4.
	c/c/d/d	3,3	1,4.	2,2	.4,1
	c/d/c/d	3,3	.2,2	.3,3.	.4,1
S_1	*d/c/c/d*	.4,1	1,4.	.3,3	.4,1
	c/d/d/c	3,3	.2,2	2,2	1,4.
	d/c/d/c	.4,1	1,4.	2,2	1,4.
	d/d/c/c	.4,1	.2,2	.3,3	1,4.
	c/d/d/d	3,3.	.2,2	2,2	.4,1
	d/c/d/d	.4,1	1,4.	2,2	.4,1
	d/d/c/d	.4,1	.2,2	.3,3.	.4,1
	d/d/d/c	.4,1	.2,2	2,2	1,4.
	d/d/d/d	.4,1	.2,2.	2,2.	.4,1

Besides the equilibrium (d,d) from the basic game (and the 2-metagame) the 1,2-metagame contains still other equilibria. It can be seen in Table 4.3 that in the 1,2-metagame (c,c) is a stable outcome in the sense that no player can improve his payoff by unilaterally changing his strategy.

Instead of the 1,2-metagame one could have taken the 2,1-metagame. This would have yielded the same equilibria. Still higher-level metagames can be derived, but Howard (1966b) has proved that the equilibria in every complete prime metagame (that is a metagame in which each player is named in the title exactly once) are the only equilibria in all higher-level metagames based on it.

Metastrategies are conditional strategies, with which a model can be build for the interaction between subjects in game-like conflict-situations. Metastrategies are no real choice-alternatives and the subjects in the game not even need to be aware of using them.

Metastrategies such as "choose c regardless of what the other does" and "choose c or d, if you think the other chooses c or d, resp. "reflect a player's orientation toward the interaction (and consequently toward the other player) and the social context in which the actual behavior takes place. Hence, it is Howard's merit that with the introduction of the concept of "metastrategies" it becomes possible to incorporate player's social orientations or motivations into a formal theory on game-playing behavior. Metagame theory is based on preference orders over the outcomes rather than the basic numerical utilities. According to the theory of metagames individual choices in a game are not based on some rationality principle, as for instance minimax (see chapter 1.), which involves a complex aggregation of utilities of different players. In other words, metagame theory avoids the alleged interpersonal incomparability of utility (see for example Arrow, 1963 and Bezembinder and Van Acker, 1979). It is a descriptive (and predictive) rather than a normative theory, since it tells us which outcomes will be stable. It does not tell us which of the stable outcomes will actually occur, but it indicates which outcomes of a game can be possibly stable (and which cannot), that is the stable outcomes in the metagames based on the basic game.

In this section a brief overview of the basic ideas of metagame theory is given for the two-person case. For more details and the n-person case the reader is referred to Howard (1971).

Let $G = (A_1, A_2; M_1, M_2)$ be a two-person game in normal form, where A_1, is the set of strategies with elements a_1 for player S_1 and M_1 is his preference function defined on the prodect-set $A = A_1 \times A_2$.

An outcome $a \in A$ of the game is defined as the joint strategy-choice $a = (a_1, a_2)$ of S_1 and S_2.

For simplicity we shall indicate player S_i as i.

$M_i : A \rightarrow B(A)$ is a set-valued function (where $B(A)$ is the set of all subsets of A), which represents i's preference as follows:

"$a \in M_i b$" (where $a, b \in A$) is read "a is not preferred by i to b".

This general form of a preference function covers the case of numerical utilities as well as other possibilities (Howard, 1974). It is assumed that M_i is reflexive, that is $a \in M_i a$ for all $a \in A$.

According to the definitions of a rational decision in Section 4.1 an outcome $\bar{a} \in A$ is rational for player 1 if and only if for all $a_1 \in A_1$:

$$(a_1, \bar{a}_2) \in M_1 \bar{a}$$

In other words, an outcome is rational for S_1, if he, S_1, cannot improve it by choosing a different strategy, given S_2's strategy choice.

$$R_1(G) = \{ \bar{a} \mid \forall a_1 : (a_1, \bar{a}_2) \in M_1 \bar{a} \}$$

$R_1(G)$ and $R_2(G)$ are non-empty for finite ordinal games (i.e. A_1 and A_2 finite and M_1 and M_2 ordinal) (Howard, 1970b, p. 214).

In the PDG of Table 4.1 $R_1 = \{(dc), (dd)\}$ and $R_2 = \{(cd), (dd)\}$. An equilibrium is an outcome which is rational for both players. The set of equilibria in G is thus:

$$E(G) = R_1(G) \cap R_2(G).$$

For the PDG of Table 4.1 this yields:

$E = \{(dd)\}$.

In experiments on game-playing behavior it is frequently observed that players attempt to predict which strategy the other player will choose. Of course, they do this in order to plan their own strategies. Player 2 choosing his strategy as if he knows what the other player is going to do or already has done, chooses (wittingly or unwittingly) from a number of conditional strategies. Such a conditional strategy or meta-strategy is a complete action-plan contingent upon the behavior of the other player. A formalized model of this situation is the special metagame $2G$.

In the special metagame $2G$ the strategy set A_2 is replaced by the set F of all fuctions f: $A_1 \to A_2$. An outcome in $2G$ corresponds to an outcome $(a_1, f(a_1))$ in G.

If G is the Prisoner's Dilemma Game (Table 4.1), then in the metagame $2G$ the outcome $(d, d/d)$ is the only equilibrium (Table 4.2). Thus (dd), the basic outcome in G which corresponds to $(d, d/d)$ in $2G$, is still the only equilibrium of the game. We say, that (dd) is the only metaequilibrium from $2G$. Metaequilibria are outcomes in the basic game which can be derived from equilibria in some metagame from G.

The metaequilibrium from $2G$ has added nothing new to what was already derived from the basic game. Nothing is gained because, of course, player's preferences for metagame outcomes are the same as for the corresponding outcomes in the basic game G.

However, something is gained in terms of rational outcomes. A meta-rational outcome for i from $2G$ is an outcome yielded by a rational outcome for i in the metagame $2G$.

From Table 4.2 we see that

$$R_2(2G) = \{(c,\ d/d),\ (c,\ d/c),\ (d,\ d/c),\ \ (d,\ c/d)\}$$

$$\approx \{(cd),\ (dd),\ (cc)\} = R_2\ (G)$$

but

$$R_1(2G) = \{(d,\ c/c),\ (d,\ d/d),\ (c,\ c/d),\ (d,\ d/c)\}$$

$$\approx \{(dc),\ (dd),\ (cc)\} \supset R_1(G)$$

It is found that (cc) is metarational for player 1 from the 2-metagame. If player 2 plays "tit-for-tat", i.e. c/d, c is the best choice for player 1. Unfortunately, there is no reason for player 2 to play "tit-for-tat".

To solve this problem Howard proposed to look at the (special) metagame $1,2G$ (Table 4.3). In $1,2G$ each metastrategy for player 1 is a function h from all $f \subset F$ of player 2 in $2G$ to A_1. An outcome in $1,2G$ is a (h,f) and the basic outcome yielded by it is obtained as

$$(h,\ f) \Rightarrow (h(f),\ f) \quad (h(f),\ f(h(f)))$$

In the 1,2-metagame of Prisoner's Dilemma it is found (see Table 4.3) that

$$E(1,2G) = \{(d/d/d/d,\ d/d),\ (d/d/c/d,\ c/d),\ (c/d/c/d,\ c/d)\}$$

which yields for the basic game: $\{(dd),\ (cc)\}$
Thus, we see that (dd) and (cc) are metaequilibria from $1,2G$ '

For example, the outcome ($c/d/c/d$, c/d), which is an equilibrium yielding (cc), involves the policy for player 1 to play c if player 2 plays c regardless or "tit-for-tat" and to play d otherwise. Player 2 plays "tit-for-tat", so player 1's strategy is c. The result is (cc).

In this way metagame theory can predict stable outcomes in experimental games; not in the sense of a normative theory, but in the sense of a descriptive theory: indicating the possible stable outcomes. Stable outcomes are defined here as outcomes resulting from each player choosing his strategy on the assumption that the other player will choose the strategy he (S_1) expects him (S_2) to choose (Howard, 1971, p. 50-51).

4.4. *General Metagames*

The metagame has been received very positively by some authors (Rapoport, 1967; Burns and Meeker, 1973). Even in the applied field it has already been proven to be a useful tool (De Beus, 1978a, 1978b; Bain, Howard and Saaty, 1971; Howard, 1968, 1972). However, others have shown themselves reserved (Shubik, 1970), or have even rejected the theory (Robinson, 1975).

Anyhow, Howard's theory has called for a number of questions as is shown for instance by the discussions in Psychological Reports (Harris, 1969a, 1969b, 1970; Rapoport, 1969, 1970; Howard, 1969, 1970a).

The main criticisms concern the asymmetric form of metagames, that is players are not treated "equally" in a metagame, and, secondly, that the prediction of the other's behavior, although one of the basic assumptions of the theory, is not provided for. Howard (1974, 1975) has solved these problems by extending the metagame concept to "general" metagames.

Let P_1 and P_2 be partitions of A_1 and A_2, respectively, in non-empty subsets indicated as p_1 and p_2, such that $\cup p_1 = A_1$ and $\cup p_2 = A_2$ and $P = (P_1, P_2)$.

Notice that $\cap p_1$ and $\cap p_2$ are not necessarily empty. If $\cap p_i = \emptyset$ than P_i is called a *strict partition.*

The game PG is played as follows:

1. Stage I : each player i (i = 1, 2) chooses a commitment $p_i \in P_i$,

2. Stage II: each player i (i = 1,2), having been informed of the other player's choice at Stage I, chooses a strategy policy $a_i \in p_i$, (i.e. from the set p_i he himself chose at Stage I), in response to the other player's commitment p_j.

By this generalization a metagame can be represented as a symmetric game in normal form and 1's (2's) choice of a strategy a_1 (a_2) is now through the choice of a p_1 (p_2) in Stage I at least partially predictable. If $p_1 = a_1$, then a_1 is entirely predictable. At the other extreme, if $P_1 = \{p_1\} = \{ A_1 \}$, a_1 is entirely unpredictable.

The special metagame $2G$ from G, discussed before turns out to be a special case. In $2G$ player 2 is assigned the partition containing, as its sole element, his total strategy-set. To player 1 is assigned the strict partition consisting of the set of all singleton sets, each of which contains just one of his strategies. In Stage II, player 2 is informed of 1's strategy, while his (2's) behavior is entirely unpredictable. This is exactly the situation that we have in the special metagame $2G$. A general metagame represents the capability of each player of predicting and enabling himself to be predicted (Howard, 1975, p.45). The subsets P_1 and P_2 which are chosen by player 1 and player 2, resp., at Stage I, are called *commitments*.

A particular general metagame is the full metagame. In the full metagame based on G the partition of each player's strategy set consists of the total set of (non-empty) subsets of his strategy-set. (If A_1 contains m elements, this partition of A_1 has 2^m-1 elements!) In the full metagame each player can make any commitment.

As an example of a full metagame Table 4.4 gives the full metagame based on the Prisoner's Dilemma in partition form. In this game the rows and columns of the original game are repeated enough times to represent the subsets belonging to each player's partition of his strategy-set.

Table 4.4 Full metagame of Prisoner's Dilemma in partition form. Each player first chooses a commitment. Then, within his commitment, each player chooses a metastrategy such that

1. *exactly one strategy is chosen to each commitment of the other player,*
2. all strategies chosen belong to a single commitment.

Examples of metastrategies are indicated:
circled cells show a metastrategy of player 1 and squared cells show a metastrategy of player 2.

		Player 2's commitments.			
		c	*d*	*c*	*d*
Player 1's commitments	*c*	3,3	1,4	$\boxed{3,3}$	1,4
	d	4,1	2,2	4,1	$\boxed{2,2}$
	c	3,3	1,4	3,3	$\boxed{1,4}$
	d	4,1	2,2	4,1	$\boxed{2,2}$

A complete metastrategy for player 1 is a pair (p_1 , f_1), where p_1 is players 1's commitment and f_1 is a function from P_2 to p_1. This function f_1, is called by Howard (1974) a *policy*. This is made clear in Table 4.5, where the full metagame in normal form is shown and where each *metastrategy* is given as a (commitment, policy)-pair.
As appears from table 4.5 only *(dd)* is a metaequilibrium from the full metagame and *(cc)* is not. However, under the metastrategy of player 2 indicated by the squared cells in Table 4.5, *(cc)* is rational for player 1. Similarly, *(cc)* is rational for player 2 under the metastrategy of player 1 indicated by the circled cells in Table 4.5, be it in a different place.

The full full metagame is the full metagame of the full metagame. In stage 1 of the full full metagame each player chooses a metacommitment, which is a subset of his set of metastrategies in the full metagame.

Table 4.5 Full metagame of Prisoner's Dilemma in normal form. Each row and each column is a complete metastrategy consisting of a commitment (the letter(s) between braces) and a policy (the triple of letters). The circled and squared cells correspond to the same metastrategies as those indicated in Table 4.4. Equilibria are underlined.

Player 2's metastrategies

Player 1's metastrategies		{*c*}	{*d*}	{*c d*}							
		c c c	*d d d*	*c c c*	*c c d*	*c d c*	*d c c*	*c d d*	*d c d*	*d d c*	*d d d*
{*c*}	*ccc*	3,3	1,4	3,3	3,3	3,3	1,4	$\boxed{3,3}$	1,4	1,4	1,4
{*d*}	*ddd*	4,1	$\underline{2,2}$	4,1	4,1	2,2	4,1	$\boxed{2,2}$	4,1	2,2	$\underline{2,2}$
{*c d*}	*ccc*	3,3	1,4	3,3	1,4	3,3	3,3	$\boxed{1,4}$	1,4	3,3	1,4
	ccd	3,3	1,4	4,1	2,2	4,1	4,1	$\boxed{2,2}$	2,2	4,1	2,2
	cdc	3,3	2,2	3,3	1,4	3,3	3,3	$\boxed{1,4}$	1,4	3,3	1,4
	dcc	4,1	1,4	3,3	1,4	3,3	3,3	$\boxed{1,4}$	1,4	3,3	1,4
	cdd	3,3	2,2	4,1	2,2	4,1	4,1	$\boxed{2,2}$	2,2	4,1	2,2
	dcd	4,1	1,4	4,1	2,2	4,1	4,1	$\boxed{2,2}$	2,2	4,1	2,2
	ddc	4,1	2,2	3,3	1,4	3,3	3,3	$\boxed{1,4}$	1,4	3,3	1,4
	ddd	4,1	$\underline{2,2}$	4,1	2,2	4,1	4,1	$\boxed{2,2}$	2,2	4,1	$\underline{2,2}$

A metacommitment of player 1 can be seen as a constraint of his own choice of (p_1 , f_1). This "constraint" can be made clear through a statement in situations of open bargaining or through a player's behavior on previous trials in games with repeated plays. Metacommitments represent a higher level of predictability than commitments.

In Stage II of the full full metagame player 1, having been informed of the "constraint" (metacommitment) chosen by player 2 in Stage I, chooses a (p_1 , f_1) from within the constraint he himself chose in Stage 1. Similarly, in Stage II player 2, having been informed of the metacommitment chosen by player 1, chooses a (p_2, f_2) from his metacommitment.

Even for a 2 x 2-game the full full metagame in partition form contains too many rows and columns to show in a figure. If H is a two-person (meta)-game with m elements in each player's (meta-strategy set, the number of commitments in the full metagame based on H is

$$\sum_{k=1}^{m} \binom{m}{k} = 2^m - 1, \qquad \text{(k is the size of the commitment)}$$

which can be represented in $\sum_{k=1}^{m} \binom{m}{k} * k = m * 2^{(m-1)}$ rows (columns) in the full metagame in partition form.

The number of metastrategies in the full metagame, based on H is

$$\sum_{k=1}^{m} \binom{m}{k} * k^{(2^m - 1)},$$

since $\underline{w}|\ \underline{y}|\ \underline{z}|\ldots$ can be filled in in $k^{(2^m-1)}$ different ways.

Table 4.6 shows how these numbers "explode" as m increases.

Table 4.6 Table representing the number of (meta)commitments, the number of rows (column) in partition form and the number of meta-(meta)strategies of the full metagame based on two-person (meta) games with different numbers of elements in the (meta)strategy sets of each player.

m = number of (meta)strategies in H	2	3	4	...	10
$2^m - 1$ = number of (meta)commitments	3	7	15	...	1023
$m * 2^{m-1}$ = number of (columns) in partition form	4	21	32	...	5120
$\sum_{k=1}^{m} \binom{m}{k}\ k^{(2^m-1)}$ = number of meta-(meta)strategies	10	2574	1,131,334,064	...	****

An important theorem concerning the identification of metaequilibria in any game is proved in Howard (1975). Since this theorem is of central importance in the experiments reported hereafter, a brief overview of the derivation will be given for the two-person case.

Let the full^k metagame of a given two-person game G be the full metagame of the full metagame of the ... full metagame of G, where the word "full" occurs k times. Then the theorem says, that for any two-person game G and any $k \geqslant 2$ the metaequilibria derived from the full^k metagame of G are the sanctionable outcomes of G. An outcome $\bar{a}$ in G is *sanctionable*, if for any (i.e. each) player there exists a strategy of the other player (a "sanction") that, if implemented, would guarantee that the former player could not do better than he does at $\bar{a}$.

In the PDG, (cc) is sanctionable. Each player has the strategy d as a sanction against the other. Also, (dd) is sanctionable, but no other outcome. If an outcome is not sanctionable in G, it is not an equilibrium in G. This is proved in the Appendix of this chapter.

Before the proof of the identification theorem is given some definitions must be made.

Definition 1 — An i-commitment is a commitment of each player i (i = 1, 2) in the full metagame. A commitment p_i of player i is a subset of his set of choice-alternatives in the basic game: $p_i \subseteq A_i$. The set of all p_i's is indicated as P_i.

Definition 2 — An i-metastrategy is a strategy of each player i (i = 1, 2) in the full metagame; it is a pair (p_1, f_1) where p_1 is an i-commitment and f_1, called an i-policy, is a function from the set of the other player's possible commitments to the basic strategies of i belonging to his chosen commitment p_i
For example: $f_1 : (p_2) \rightarrow a_1 \in p_1$.
The set of all f_i's for a given p_i is indicated as $F_i(p_i)$

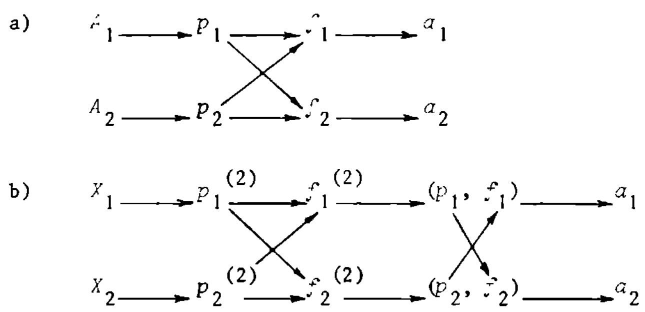
The following definitions relate to the full full metagame, referred to as the full^2 metagame.

Definition 3 An i-metacommitment $p_i(2)$ is an i-commitment in the full^2 metagame: $p_i(2) \in X_i$ with $X_i = \{(p_i , f_i) \mid p_i \in P_i ; f_i \in F_i (p_i)\}$

Definition 4 An i-metastrategy in the full^2 metagame is a pair $(p_i^{(2)} , f_i^{(2)})$, where $f_i^{(2)}$ is a function from the set of the other player's possible metacommitments to the metastrategies of i belonging to his chosen metacommitment $p_i^{(2)}$.

Figure 4.1 gives a tree diagram for the decision making process in the full metagame and the full^2 metagame.

Figure 4.1



Tree diagram for the decision making process in the full metagame (a) and the full^2 metagame (b).

A metaequilibrium from a given metagame h is an outcome in the basic game, which is derived from an equilibrium in H

Let $[(\bar{p}_1, \bar{f}_1) , (\bar{p}_2, \bar{f}_2)]$ be an equilibrium of the full metagame, and let $\bar{f}_1(\bar{p}_2) = \bar{a}_1$ and $\bar{f}_2 (\bar{p}_1) = \bar{a}_2$. Then $\bar{a} = (\bar{a}_1, \bar{a}_2)$ is a metaequilibrium from the full metagame.

The following lemma is proved in Howard (1974).

Lemma 1. If $\bar{a} = (\bar{a}_1, \bar{a}_2)$ is a metaequilibrium from a metagame H, it is a metaequilibrium from every further metagame based on H.

Another important lemma is proved in Howard (1975).

Lemma 2. If $\bar{}$ is not sanctionable in G, no metagame-outcome $[(\bar{p}_1, \bar{f}_1), (\bar{p}_2, \bar{f}_2)]$ that yields $\bar{a}$ is sanctionable in the full metagame of G.

Since "not sanctionable" implies "not an equilibrium", this lemma says that only the sanctionable outcomes can be metaequilibria from the full^2 metagame.
For the special metagames it was demonstrated (Howard, 1971) that the sanctionable outcomes of any game G are the symmetric equilibria of G, that is outcomes that are metarational for all players in some derived metagame from any metagame based on G.

After lemmas 1 and 2 the identification of metaequilibria can be solved, if it is proved that all sanctionable outcomes in the basic game are metaequilibria from the full^2 metagame. This proof is given in Howard (1975).

Theorem If $\bar{a} = (\bar{a}_1, \bar{a}_2)$ is sanctionable in G, $\bar{a}$ is a metaequilibrium from the full^2 metagame of G.

To prove this theorem Howard (1975, p. 19) introduced the $\bar{a}$-enforcing i-metacommitment in the full^2 metagame:

$$\{p_i, f_i \mid p_i = \{\bar{a}_i\} \text{ or } A_i; \text{ but if } p_i = A_i, f_i(p_j) = \bar{a}_i, \text{ when } p_j = \{\bar{a}_j\}\}$$

Under the $\bar{a}$-enforcing i-commitment, i's commitment is either $\{\bar{a}_i\}$ or A_i; if i makes the latter commitment, at least he makes the constraint policy to choose $\bar{a}_i$, if the other player commits himself to $\{\bar{a}_j\}$.

The $\bar{a}$-enforcing i-commitment models a choice behavior with the following features:

1. if player j makes the $\bar{a}$-enforcing j-metacommitment, i makes the commitment $\{\bar{a}_i\}$;
2. if player j does not make the $\bar{a}$-enforcing j-metacommitment, i makes the commitment A_i and chooses within this the strategy a_i' which results in an outcome a' that is less preferred by j than $\bar{a}$, that is $a' \in M_j\bar{a}$.

This latter policy is called an $\bar{a}$-enforcing i-metapolicy.

Now it follows that if each player i chooses a metastrategy $(p_i^{(2)} f_i^{(2)})$, where $p_i^{(2)}$ is the $\bar{a}$-enforcing i-metacommitment and $f_i^{(2)}$ is an $\bar{a}$-enforcing i-metapolicy, the result is an equilibrium of the full^2 metagame that yields $\bar{a}$.

4.5. *A test of Howard's Theory.*

Experimental tests on the theory of "special games" have already been performed by a number of researchers. See for example Howard (1970), Thomas (1973, 1974) and Hildebrand (1973).
A detailed description on how to design and conduct (special) metagame-experiments is given by Thomas (1974). All of these experiments apply to two person games where the strategy of the opponent is fully predictable. So far no experiments on general metagames are known.

The theory of metagames is a theory on stable outcomes in games. It is based on some assumptions about the decision-making process which precedes the actual choices of the players in the game. A stable outcome is an outcome anticipated by each player, i.e. an outcome that is determined by each player's individual choice and each player's individual prediction.

To test the theory on general metagames an "extended" Prisoner's Dilemma Game (EPDG) was constructed. This 4*4-game is identical to a usual PDG except that the basic outcomes of PDG are replaced with 2*2-sub-games (see Table 4.7).

Table 4.7

		S_2			
		c		*d*	
		w	*x*	*y*	
S_1 *c*	*w*	14,14	10,13	5,15	1,16
	x	13,10	9,9	7,11	3,12
d	*y*	15,5	11,7	6,6	2,8
	z	16,1	12,3	8,2	4,4

"Extended" Prisoner's Dilemma Game. The number pairs in each cell indicate ordinal preferences - higher 'pay-offs' being more preferred.

Each player's strategy set in this game consists of four elements, w, x, y, and z, with y and z dominant over w and x, respectively. In the "rational" quadrant, which is formed by the strategies y and z of both players, each outcome is less preferred than any outcome in the "cooperative" quadrant, which is formed by the strategies w and x of both players. Each subgame, in which y or z is chosen by at least one player, is a usual PDG.

If the game is played as a 4*4-game (i.e. each player chooses one strategy directly from his strategy set), there are 10 sanctionable outcomes (the circled outcomes in Table 4.7). However, if the game is played in two stages (i.e. first choose a subgame and then choose a strategy in this subgame), the picture changes.

The major problem in designing an experiment on general metagames is to determine empirically the (non-observable) metastrategies of the players. "Choosing" a policy from a one-element commitment in the full metagame of PDG may seem absurd, when doing it consciously. However, choosing a policy from a two-element commitment is not. Playing EPDG as described in the last paragraph is similar to playing in the full metagame of PDG, when c is replaced with w and x, and d is replaced with y and z (Table 4.7). Considering the EPDG as a PDG with subgames as the basic outcomes, it is seen that the "rational" quadrant is an equilibrium subgame and the "cooperative" quadrant is a sanctionable subgame. Through the EPDG we are able to have subjects play the full metagame of the PDG without them being aware of it.

4.5.1. *Hypotheses.*

1. According to Howard the stable outcomes, that is the outcomes the players will agree upon, are sanctionable outcomes.

   For the PDG this means, that (cc) and (dd) are possible stable outcomes.

2. A sanctionable outcome $\bar{a}$ is attained by each player i choosing an element from the $\bar{a}$-enforcing i-metacommitment. In the PDG the (cc)-enforcing i-metacommitment consists of the following metastrategies:

   - $\{c\}$, $c/c/c$
   - $\{c, d\}$, $c/d/c$
   - $\{c, d\}$, $c/d/d$

In the latter two metastrategies player i chooses c if the other player commits himself to $\{ c \}$ and i chooses d if the other player commits himself to $\{ d \}$. Thus, when player i makes the "empty" commitment, i.e. $\{ c,d \}$, he promises to use his sanction against the other player j for not making the (cc)-enforcing j-metacommitment. By the same line of reasoning we find that the (dd)-enforcing i-metacommitment consists of the metastrategies:

- $\{d\}$, $d/d/d$
- $\{c\ , d\}$, $d/d/c$
- $\{c\ , d\}$, $d/d/d$.

4.5.2. *Experimental procedure*

The theory of metagames asserts that a certain prediction will be confirmed in all cases: a single substantiated counterexample suffices to overthrow the theory (Thomas, 1974, p. 75). As a consequence, metagame experiments are completely deterministic in nature. By this we mean that the aim of metagame experiments is to obtain a falsification of the theory, if such falsification can be found. In contrast, probabilistic or statistical experiments are concerned with seeking probabilistic estimates of the validity of a theory. One of the main principles in designing deterministic experiments is to find a procedure that evokes the behavior about which assertions are being made.

Seven groups of 4 subject in each participated in the experiment. These 28 subjects were all undergraduate university students.

The four subjects in each group are denoted I, J, K, and M. First instructions were given to I and J by one experimenter and to K and M by a second experimenter. During the instruction I and J were kept separated from K and M.

To ensure that the subjects understood the game they played a practice game against each other (i.e. I against J, and K against M), till both they and the experimenter were satisfied on that count. During practising the players could communicate freely with the experimenter about the rules of the game.

The practice game as well as the actual game were similar to the game shown in Table 4.7 in that the outcomes were assigned monetary payoffs in the same order of magnitude as the numbers in the EPDG of Table 4.7. This was to make sure that the players had the same preferences as assigned to the outcomes in the EPDG. The monetary payoffs differed for the practice game and the actual game. The payoff matrix of the actual game is displayed in Table 4.8.
After instruction the actual game was played against a different subject, that is I played against K and J played against M. Opponents in

the actual game did not know each other. They were kept separated from each other and were not allowed to communicate with each other. They were told that their objective should be to make money for themselves. Earnings from the actual game would be payed.

Table 4.8 Payoff matrix of the experiment. The payoffs are in Dutch Guilders and cents.
The players of the game were indicated as "Red" and "Blue".

		Blue			
		W	X	Y	Z
Red	W	8.80 / 8.80	6.40 / 8.20	3.40 / 9.40	1.00 / 10.00
	X	8.20 / 6.40	5.80 / 5.80	4.60 / 7.00	2.20 / 7.60
	Y	9.40 / 3.40	7.00 / 4.60	4.00 / 4.00	1.60 / 5.20
	Z	10.00 / 1.00	7.60 / 2.20	5.20 / 1.60	2.80 / 2.80

A procedure was set up to evoke the behavior about which assertions are made in the theory of general metagames and which were to be tested. As outlined above the EPDG of Table 4.8 was played such as to make possible an experimental determination of the metastrategies of the players.

The game was played in two stages:

1. choose a subset from the set of choice alternative,
2. select from this subset a strategy to each possible subset of the other player.

The following instruction was given to the subjects.

"To help you in making your choice of an alternative we shall proceed in two steps. Your choice will be determined by answering two questions.

Although you can make a choice from 4 alternatives, w, x, y and z, it may be possible, that you *exclude one or two alternatives in advance.*

<u>Question 1.</u> *From which alternatives are you going to choose a strategy?*

- ☐ *from w and x only*
- ☐ *from y and z only*
- ☐ *from w, x, y and z.*

Answer question 2 only, after you answered question 1!

Question 2. *From the alternatives indicated under question 1 I choose*

--- *if the other subject chooses a strategy from w and x only*

--- *if the other subject chooses a strategy from y and z only*

--- *if the other subject chooses a strategy from w, x, y and z.*

The answer to question 1 determines a commitment, and the answer to question 2 is a policy. The (commitment, policy) pairs, or metastrategies, of both players yield together an outcome of the game.

Subjects wrote down their answers, that is their metastrategies, without knowing wath the other player had chosen or was going to choose.

After the experimenter had received the notes containing each player's metastrategy he announced the commitments made, that is the subsets chosen under question 1, and the outcome resulting from the (commitment, policy) pairs of the players. The experimenter then asked each player to indicate in writing whether he (i.e. the player) would like to end the game and being paid according to the payoff of the last outcome or whether he would like to continue the game and try to improve his payoff. As long as at least one player wanted to continue the game the whole procedure as described above was repeated.

When both players decided to end the game, they were paid according to the payoff of the last outcome.

4.5.3. *Results*

Table 4.9 shows the "raw" stable outcomes of 26 subjects, that is the outcomes in the EPDG which were attained when both players decided to end the game. The results of one pair of subjects are omitted, because they did nót adhere to the instructions of the experimenter and communicated with each other.

Table 4.9 Stable outcomes in the Extended Prisoner's Dilemma Game for 13 pairs of subjects. "$w/z/w$" means: choose w when the other player makes the commitment $\{w,x\}$, choose z when the other player makes the commitment $\{y,z\}$, and choose w when the other player makes the commitment $\{w,x,y,z\}$.

pair of subjects	Red's metastrategy commitment	policy	Blue's metastrategy commitment	policy	Outcome
1	$\{w,x,y,z\}$	$w/z/w$	$\{w,x\}$	$w/w/w$	(ww)
2	$\{w,x\}$	$x/x/x$	$\{w,x,y,z\}$	$x/z/z$	(xx)
3	$\{w,x\}$	$w/x/w$	$\{w,x,y,z\}$	$w/y/w$	(ww)
4	$\{w,x\}$	$w/w/w$	$\{w,x,y,z\}$	$x/z/z$	(wx)
5	$\{w,x,y,z\}$	$x/z/w$	$\{w,x\}$	$w/x/w$	(xw)
6	$\{w,x\}$	$w/x/w$	$\{w,x\}$	$w/x/w$	(ww)
7	$\{w,x\}$	$w/x/w$	$\{w,x,y,z\}$	$w/y/w$	(ww)
8	$\{w,x,y,z\}$	$z/z/w$	$\{y,z\}$	$y/z/z$	(zz)
9	$\{w,x,y,z\}$	$w/y/w$	$\{w,x\}$	$w/x/w$	(ww)
10	$\{w,x\}$	$w/x/w$	$\{w,x,y,z\}$	$w/z/w$	(ww)
11	$\{w,x\}$	$w/x/w$	$\{w,x,y,z\}$	$w/z/w$	(ww)
12	$\{w,x,y,z\}$	$w/z/x$	$\{w,x\}$	$w/x/w$	(ww)
13	$\{w,x,y,z\}$	$w/z/w$	$\{w,x\}$	$w/x/w$	(ww)

It was found that 12 out of 13 pairs of players "settled" in the "cooperative" quadrant.

The commitments the players could make were $\{w, x\}$, $\{y, z\}$ and $\{w, x, y, z\}$. These are not all possible commitments in a full metagame. However, if we "lump" w and x to c, and y and z to d, we have all commitments in the full metagame of a Prisoner's Dilemma Game: $\{c\}$, $\{d\}$, $\{c,d\}$.

Table 4.10 shows the experimental results after this lumping.

From Table 4.10 it can be seen, that (cc) and (dd), the sanctionable outcomes in PDG, are the stable outcomes in the game.

Table 4.10 Stable outcomes with the Extended PDG after "lumping" w and x into c, and y and z into d.

Pair of subjects	Red's metastrategy commitment	policy	Blue's metastrategy commitment	policy	Outcome
1	$\{c,d\}$	$c/d/c$	$\{c\}$	$c/c/c$	(cc)
2	$\{c\}$	$c/c/c$	$\{c,d\}$	$c/d/d$	(cc)
3	$\{c\}$	$c/c/c$	$\{c,d\}$	$c/d/c$	(cc)
4	$\{c\}$	$c/c/c$	$\{c,d\}$	$c/d/d$	(cc)
5	$\{c,d\}$	$c/d/c$	$\{c\}$	$c/c/c$	(cc)
6	$\{c\}$	$c/c/c$	$\{c\}$	$c/c/c$	(cc)
7	$\{c\}$	$c/c/c$	$\{c,d\}$	$c/d/d$	(cc)
8	$\{c,d\}$	$d/d/c$	$\{d\}$	$d/d/d$	(dd)
9	$\{c,d\}$	$c/d/c$	$\{c\}$	$c/c/c$	(cc)
10	$\{c\}$	$c/c/c$	$\{c,d\}$	$c/d/c$	(cc)
11	$\{c\}$	$c/c/c$	$\{c,d\}$	$c/d/c$	(cc)
12	$\{c,d\}$	$c/d/c$	$\{c\}$	$c/c/c$	(cc)
13	$\{c,d\}$	$c/d/c$	$\{c\}$	$c/c/c$	(cc)

It is interesting to see under which metastrategies subjects arrived at (cc) and then decided to end the game, that is under which metastrategies stability was attained at (cc).
In Table 4.11 a summary of the metastrategies which were chosen when stability was attained at (cc), is given.

Table 4.11 Summary of the metastrategies under the (cc)-enforcing metacommitments when stability was reached.

commitment	policy	frequency
$\{c\}$	$c/c/c$	13
$\{c,d\}$	$c/d/c$	9
$\{c,d\}$	$c/d/d$	2
		n=24

It appears that all subjects, when reaching stability at (cc), had chosen an element from the (*cc*)-enforcing i-metacommitment, although there are more metastrategies which can yield (*cc*) (see Table 4.5).

One pair of subjects settled at (*dd*). They continued the game for ten trials before both players decided to end the game (the average number of trials for the other pairs of subjects was 3.25). "Red" set out with the (commitment, policy)-pair [{*c*, *d*}, *c*/*d*/*c*] and "Blue" with [{*c*, *d*}, *d*/*d*/*d*]. This yielded (*cd*). Then, seven trials followed where Red continously chose [{*c*}, *c*/*c*/*c*] and Blue chose [{*d*}, *d*/*d*/*d*] or [{*c*, *d*}, *d*/*d*/*d*]. So, all the the time the outcome of the game was (*cd*), which was unsatisfying to Red, who was choosing an element from the (*cc*)-enforcing metacommitment, while Blue persisted in obviously the (*dd*)-enforcing metacommitment.

On the nineth trial Blue switched to [{*cd*}, *c*/*d*/*d*], and since Red still chose [{*c*}, *c*/*c*/*c*], this yielded (*cc*).
This outcome satisfied Red, but now it was Blue, who did not want to end the game. On the tenth trial Blue returned to the "rational" metastrategy [{*d*}, *d*/*d*/*d*] and Red, clearly getting annoyed, chose [{*c*, *d*}, *d*/*d*/*c*]. Then both players decided to end the game.

4.6 *Discussion*

The theory of metagames is a theory about stable outcomes. The predictions made by general metagame theory were not falsified in our experiment. First, all pairs of subjects, reached stability at a sanctionable outcome and, second, sanctionable outcomes $\bar{a}$ were attained by each player i choosing an element from the $\bar{a}$-enforcing i-metacommitment.

It is not very surprising that most pairs of subjects settled at (*cc*), since the game was continued until both subjects were satisfied. As such, these "games with certainty" were not very informative. More important are the metastrategies through which stability was attained, that is through the subjects choosing elements from the (*cc*)-enforcing metacommitment and the (*dd*)-enforcing metacommitment respectively

The significance of metagame theory does not lie in its power to predict the stable outcomes, but in the way the decision making process in modelled as a cognitive process of beliefs (predictions) and choosing policies.

Discrepancies which are so often found in nonzero-sum games, like PDG, between actual behavior and the prescriptions of classical game theory, are deviations from rational behavior, i.e. by introducing metarationality, Howard has formalized these kinds of "deviant" behavior, which are often labeled with attitudinal epithets as cooperation, altruism, competitiveness, etcetera. Cooperative behavior in the Prisoner's Dilemma Game is no longer irrational behavior.

In iterated plays of a game stability of behavior is often found when subjects absorb or "lock-in" at certain outcomes. In iterated plays of PDG it is found (see for example Rapoport and Chammah, 1965) that most pairs of players lock-in at (dd) or (cc). Thus, in PDG subjects' behavior stabilizes or absorbs into sanctionable outcomes or metaequilibria of the game.

The experiment reported in this paper was an attempt to test some predictions of general metagame theory and to have the players made explicit their commitments and policies. So far no other experimental tests on general metagame theory are known to the author. Although the approach seems useful to determine empirically the metastrategies of the players, if has to be developed further. For example, one might wish to extend this approach to multi-person games, or to games with uncertainty to see how metastrategies, that is beliefs and policies, change in the course of interaction under risk.

Another major question concerns the exact nature of commitments and metacommitments. Although theoretically the number of metastrategies is large (see Table 4.6), it might be that the number of metastrategies actually chosen in a game, is restricted and that these metastrategies can be arranged in classes reflecting certain social orientations between the players. A similar idea was already launched by Burns and Meeker (1973), who assume that "... *social orientations operate as metaprocesses to tranform preferences, decision procedures, and interaction patterns*" (p. 159).

A final note concerns the method of experimentation. The validity of the experimental results and their interpretation stands or falls with the way the game is perceived by the players. That is, do the players understand the game and do they have the preferences that the experimenter wishes them to have?

Understanding the game means knowing the rules of the game and knowing what each player's possible choices of strategies are. To accomplish this the practice game was introduced, which was similar to the actual game in all respects but for the monetary payoff magnitudes.

Players' preferences can be contaminated in many ways. Several steps were taken to exclude this source of experimental error. First, to prevent contamination due to effects of learning and experience, players never practised with anyone who was also an opponent in the actual game. Second, to prevent contamination due to altruism, players in the actual game were isolated from each other throughout the experiment. Third, as already mentioned, outcomes were assigned monetary payoffs in the same order of magnitude as the order of preferences assigned by the experimenter. Furthermore, the magnitudes of payoffs were chosen such as to yield differences meaningful to the subjects.

The results of Table 4.10 could be criticized as the result of still another kind of preference contamination. With respect to the occurrence of (cc) in 12 out of 13 cases it could be argued that all pairs of players except one had interpreted the task as a zero-sum game with the pair of subjects as one player and the experimenter as the opponent. The objective would have been to maximize to joint gain. Two arguments can be brought in against such a critisism. First, in the instruction subjects were told to choose such as to make money for themselves. Second, from Table 4.11 it appears that about half of the subjects had chosen a metastrategy under the (cc)-enforcing meta-commitment with a policy in wich d is chosen when the opponent does not make the $\{c\}$-commitment. This is non-optimal when both players try to maximize their joint payoff. This result does not support the hypothesis of a "conspiracy against the experimenter".

Arrow, K.J. *Social choice and individual values.* New York: Wiley, 1963.

Bain, H., Howard, N. and Staaty, T.L. Using the Analysis of Options Technique to Analyze a Community Conflict. *Journal of Conflict Resolution,* 1971, *II*, 133-144.

Bezembinder, Th.G.G. and Van Acker, P.O. Intransitivity in individual and social choice. In: E.D. Lantermann and H. Feger (Eds.), *Similarity and choice.* Bern: Huber, 1979.

Burns, T. and Meeker, L.D. A Mathematical model of Multi - Dimensional Evaluation, Decision Making, and Social Interaction. In: James L. Cochrane and Milan Zeleny (Eds.), *Multiple Criteria Decision Making.* South Carolina: The University of South Carolina Press, 1973.

De Beus, J.W. De onafwendbaarheid van een geleide loonpolitiek (I). *Economische Statistische Berichten,* 1978a, *63.* 132-137 (in Dutch).

De Beus, J.W. De onafwendbaarheid van een geleide loonpolitiek (II). *Economische Statistische Berichten,* 1978b, *63.* 173-178 (in Dutch).

Friedland, N., Arnold, E., and Thibaut, J. Motivational Bases in Mixed-Motive Interactions: The effects of Comparison Levels. *Journal of Experimental Social Psychology,* 1974, *10*, 188-199.

Griesinger, D.W. and Livingston, J.W.Jr. Toward a Model of Interpersonal motivation in Experimental games. *Behavioral Science,* 1973, *18*, 173-188.

Harris, R.J. Note on Howard's Theory of Metagames. *Psychological Reports,* 1969, *24*, 849-850 (a).

Harris, R.J. Comments on Dr. Rapoport's Comments. *Psychological Reports,* 1969, *25*, p. 825 (b).

Harris, R.J. Paradox regained. *Psychological Reports,* 1970, *26*, 264-266.

Hildebrand, P. *Stable outcomes in Two-Person Games.* Ph.D. Thesis, University of Waterloo, 1973.

Howard, N. The Theory of Metagames. *General Systems,* 1966, *XI*, 167-186 (a).

Howard, N. The Mathematics of Metagames. *General systems,* 1966, *XI*, 187-200 (b).

Howard, N. Metagame Analysis of Vietnam Policy. *Papers, Peace Research Society (International),* 1968, *Vol. X.*

Howard, N. Comments on Harris' "Comments on Rapoport's Comments". *Psychological Reports,* 1969, *25*, p. 826.

Howard, N. Note on the Harris-Rapoport Controversy. *Psychological Report,* 1970, *26*, p. 316 (a).

Howard, N. Some developments in the theory and application of metagames. *General Systems,* 1970, *XV*, 205-231 (b).

Howard, N. *Paradoxes of Rationality: Theory of Metagames and Political Behavior.* Massachusetts: MIT Press, 1971.

Howard, N. The Arab-Israëli Conflict: A Metagame Analysis Done in 1970. *Papers, Peace Research Society (International),* 1972, *Vol. XIX.*

Howard, N. "General" Metagames: An Extension of the Metagame Concept. In: A Rapoport (Ed.), *Game Theory as a Theory of Conflict Resolution.* Dordrecht-Holland: Reidel Publishing Company, 1974.

Howard, N. Prisoner's Dilemma: The "Solution" by General Metagames. *University of Ottawa, Working Paper 75-24,* 1975.

Krivohlavý, J. *Zwischenmenschliche Konflikte und Experimentelle Spiele.* Bern: Verlag Hans Huber, 1974 (in German).

Luce, R.D. and Raiffa, H. *Games and decisions.* New York: Wiley and Sons, 1957.

Messick, D.M. and McClintock, C.G. Motivational bases of choice in experimental games. *Journal of Experimental Social Psychology,* 1968, *4*, 1-25.

Messick, D.M. and Thorngate, W. Relative gain maximization in experimental games. *Journal of Experimental Social Psychology,* 1967, *3*, 85-101.

Morehous, L.G. Two Motivations For Defection in Prisoner's Dilemma Games. *General Systems,* 1966, *XI,* 225-228.

Rapoport, A. Models of Conflict: cataclysmic and strategic. *C.I.B.A. Symposium on Conflicts in Society,* 1966, 259-287.

Rapoport, A. Escape from Paradox. *Scientific American,* July 1967, 50-56.

Rapoport, A. Comments on Dr. Harris' "Note on Howard's Theory of Metagames". *Psychological Reports,* 1969, *25*, 765-766.

Rapoport, A. Comments on "Paradox Regained". *Psychological Reports,* 1970, *26*, p. 272.

Rapoport, A. and Chammah, A. *Prisoner's Dilemma.* Ann Arbor: University of Michigan Press, 1965.

Robinson, M. Prisoner's Dilemma: Metagames and Other Solutions. *Behavioral Science,* 1975 *20,* 201-205.

Schelling, T.C. *The strategy of Conflict.* Cambridge: Harvard University Press, 1960.

Shubik, M. Game Theory, Behavior and the Paradox of the Prisoner's Dilemma: Three Solutions. *Journal of Conflict Resolution,* 1970, *2,* 181-193.

Terhune, K.W. Motives, situation, and interpersonal conflict within Prisoner's Dilemma. *Journal of Personality and Social Psychology,* 1968, *8*, monograph supplement.

Thomas, C.S. *Transitions, and Offered Sets in Two-Person Games.* Ph.D.Thesis, University of Waterloo, 1973.

Thomas, C.S. Design and Conduct of Metagame Theoretic Experiments. In: A. Rapoport (Ed.), *Game Theory as a Theory of Conflict Resolution.* Dordrecht-Holland: Reidel Publishing Campany, 1974.

Von Neumann, J. and Morgenstern, O. *Theory of Games and Economic Behavior.* Princeton: Princeton University Press, 1944.

We prove here, that a non-sanctionable outcome in G is not an equilibrium in G.

If M_i is the relation *not preferred by i*, then $\tilde{M}_i$ is the relation *preferred by i*. In any ordinal game the relation $\tilde{M}_i$ is transitive. (Howard, 1971, p. 71-72).

1. If $(\bar{a} = (\bar{a}_1, \bar{a}_2)$ is not an equilibrium in G, there exists for some player i a strategy a_i', with $a_i' \neq \bar{a}_i$, such that $(a_i', \bar{a}_j)$ is preferred by i to $\bar{a}$. :

   $$\exists\, a_i' : (a_i', \bar{a}_j) \in \tilde{M}_i\, \bar{a}$$

2. If $\bar{a} = (\bar{a}_1, \bar{a}_2)$ is not sanctionable in G, there exists for some player i a strategy a_i', with $a_i' \neq \bar{a}_i$, such that for any a_j the outcome (a_i', a_j) is preferred by i to $\bar{a}$.

   $$\exists a_i' \forall a_j : (a_i', a_j) \in \tilde{M}_i \bar{a} \Rightarrow \exists a_i' : (a_i', \bar{a}_j) \in \tilde{M}_i \bar{a}$$

3. From (1) and (2) it can be seen, that "not sanctionable" implies "not an equilibrium". Q.e.d.

Chapter 5: A dual Markov chain for the Prisoner's Dilemma Game, with three latent states

5.1. *Introduction*

Although several authors have stressed the importance of players' motivational orientation toward the game situation as determinants of their (i.e. the players') behavior (Morehous, 1966; McClintock and McNeel, 1966; Messick and McClintock, 1968; Ofshe and Ofshe, 1970; Pruitt, 1970; Wyer, 1971; Burns and Meeker, 1973; Kuhlman and Marchello, 1975), this has not resulted yet in process models which describe the time course of interaction as a result of the relation between latent characteristics of the players and their manifest behavior.

According to metagame theory (Howard, 1966, 1971) the sanctionable outcomes of a game can be stable (Howard, 1976). This was discussed in detail in the last chapter. Remember, that a sanctionable outcome in a two-person game is an outcome such that for any player there exists a strategy of the other player that, if implemented, would quarantee that the former player could not do better than he does at this (sanctionable) outcome. In the Prisoner's Dilemma Game (PDG) (see Table 2.1) the outcomes (cc) and (dd) are sanctionable.

According to general metagame theory (Howard, 1976) stability is attained through subjects choosing metastrategies consisting of a (commitment, policy) pair. A commitment was defined in chapter 4 as a subset of a player's strategy set which he chooses himself as a constraint. A policy was defined as a function from the possible commitments of the other player to the strategies within the own commitment. By means of the propositions of the theory of metagames possible equilibria, in the sense of stable outcomes, are derived for any game. This can also be applied to games with repeated plays. It has been pointed out already that the theory of metagames reached "solutions" which are not given by classical game theory. Experiments of Howard (1971), Thomas (1973, 1974), Hildebrand (1973) and Van der

Sanden (1979) (see also last chapter) have demonstrated that subjects in experimental games do behave according to the predictions of metagame theory.

There is a similarity between Howard's metastrategies and the concept of *value states* as applied by Meeker (1971), Leik and Meeker (1975, chapter 10) and Van der Sanden (1978; see also chapter 2). A value state and a metastrategy both have to do with the subjective principle of justice (subjective rationality or metarationality in Howard's terms) of a player with regard to his most desirable behavior.

Meeker's value conflict model has provided a way of analyzing the process of decision making of subjects in iterated plays of nonzero-sum games such as Prisoner's Dilemma. Meeker's approach was concerned in the underlying structure of interaction in a conflict and the dynamic relation between the latent behavior (a subject's motivational orientation or his value state) and the manifest behavior, i.e. observable choice behavior.

The experiments of Meeker and Van der Sanden both applied only to conflict situations, where the subject's opponent was a preprogrammed stooge. Before a "value state" model can be applied to a "free" interaction, the orginal model of Meeker should be revised. Meeker's latent Markov model will have to be expanded with absorbing states representing the states which the players occupy when their behavior becomes "absorbed" permanently, that is stabilizes, in a specific outcome of the game. The nature and the number of these states are derived from Howard's theory of metagames.

Rapoport and Dale (1966) have evaluated five models for behavior in the PDG with respect to their efficiency to describe the observed time course of outcomes in this game over a great number of trials. Three of their models were Markov chain models for the observed outcomes. It appeared that Markov chain models which are based exclusively on the observable behavior, showed the greatest discrepancies between predicted and observed time courses of their data. Rapoport and Dale gave as their comment. "... *we feel that the* [...] *models do not capture the dynamics of the process and must be discarded*" (1966, p. 285).

The two models, which were satisfying, were a Markov chain model with absorbing states, which shows great resemblance with Cohen's conflict model as applied in his conformity experiments (Cohen, 1963), and a stochastic learning model with linear operators according to Bush and Mosteller (1955) and Estes (1950) for the response parameter.

The models of Rapoport and Dale were compared on their descriptive power of observed behavior sequences of male and female pairs of players in the PDG. The latter two models suffered from the disadvantage that:

1. the number of parameters were too large, and therefore
2. a procedure for estimating the parameters was lacking.

The parameters which fitted the data best, for the various models were found through a method of trial-and-error.

The aim of this chapter is to present and to evaluate a mathematical model which describes the choice behavior of a subject in a PDG played over many plays in succession and which is based on theoretical concepts from Howard's theory of metagames and the "value state" approach proposed by Meeker.

The main characteristics of this model are:

1. the behavior of each player is described by a Markov model with states not observable over plays, a so-called latent Markov chain model,
2. the two-player system is represented as a finite probabilistic automaton,
3. the parameters of this automaton are estimated through a maximum likelihood procedure,
4. with the parameters different unobservable quantities can be computed, such as the average number of trials to absorption, the variance of this variable, the probability distribution over the final states of the automaton.

5.2. *LAMARC, a latent Markov chain model*

Two subjects, S_1 and S_2, play the PDG a great number of times in succession. After each play the outcome is announced before the next play. It is assumed that on each play n ($n = 1, 2, \ldots$) of the game each subject is in exactly one of a set of latent states. Each latent state is characterized by a subject's expectation of his own behavior and his expectation of the behavior of the other player. These latent states are called *value states*, since they represent the norms or values of the subject with regard to his most desirable behavior on te next play. A subject's responses or strategy-choices on each play are dependent upon the outcomes of the game on earlier plays and upon the value states.

Definition 1: Let A_i be the set of strategies for S_i (i = 1, 2). In the PDG these strategies are denoted as c and d. (see Table 2.1): Thus, $A_i = \{c, d\}$ for each S_i (i = 1, 2).

Definition 2: Let A be the set of outcomes of the game. A is defined as the Cartesian product of the sets of strategies. Thus, $A = A_1 \times A_2 = \{(cc), (cd), (dc), (dd)\}$.

Definition 3: Let V_i be the set of value states for S_i. It has three elements: $V_i = \{v_1, v_2, v_3\}$.
A subject in state v_1 expects the other player to choose c and also himself to choose c on the next play. It is called the "cooperative state".
A subject in state v_2 neither knows what to expect from the other player nor has committed himself to c or d for the next play. It is called the "transitional state".
A subject in state v_3 expects the other player to choose d and also himself to choose d on the next play. It is called the "rational state".

The above-mentioned definitions are the basis for the following response axioms.

<u>Response axioms</u>

R1: When S_i (i = 1, 2) is in state v_1, he will choose c on the next play with probability 1.

R2: When S_i ($i = 1, 2$) is in state v_2, he will choose c on the next play with probabliblity π_i.

R3: When S_i ($i = 1, 2$) is in state v_3, he will choose d on the next play with probability 1.

The conditional choice probabilities which follow from these axioms are summarized in Table 5.1.

Table 5.1 Conditional choice-probabilities for one subject

		strategies	
		c	d
	v_1	1.	0.
value states	v_2	π_i	$1-\pi_i$
	v_3	0.	1.

A subject enters a value state as the result of his value state on the preceding play and the outcome of the preceding play.

If a subject is in v_2 and the last outcome was asymmetric, i.e. (cd) or (dc), he will be in v_2 on the next play as well. The transition from v_2 to v_1 is possible only immediately after a double-cooperating outcome, i.e. (cc). The transition from v_2 to v_3 is possible onlv immediately after a double-defecting outcome, i.e. (dd).

If S_i ($i = 1, 2$) is in v_1, he will continue to be in v_1 as long as (cc) is chosen. Otherwise he will be in v_2 on the next play with probability β_i and in v_1 with probability $(1 - \beta_i)$.

If S_i ($i = 1, 2$) is in v_3, he will continue to be in v_3 as long as (dd) is chosen. Otherwise he will be in v_2 on the next play with probability γ_i and in v_3 with probability $(1 - \gamma_i)$.

One-step stransitions from v_1 to v_3 and from v_3 to v_1 are excluded.

Transition axioms (see Figure 5.1a):

T1a: If S_i ($i = 1, 2$) occupies v_1 on play n and both players choose c on play n, S_i will continue to be in v_1 on play n + 1 with probability 1.

T1b: If S_i ($i = 1, 2$) occupies v_1 on play n and the other player chooses d on play n, S_i will be in v_2 with probability β_i and in v_1 with probability $1 - \beta_1$ on play n + 1.

T2a: If S_i ($i = 1, 2$) occupies v_2 on play n and both players choose c on play n, S_i will be in v_1 with probability α_i and in v_2 with probability $1 - \alpha_i$ on play n + 1.

T2b: If S_i ($i = 1, 2$) occupies v_2 on play n and the outcome on play n is (cd) or (dc), S_i will continue to be in v_2 on play n + 1 with probability 1.

T2c: If S_i ($i = 1, 2$) occupies v_2 on play n and both players choose d on play n, S_i will be in v_3 with probability δ_i and in v_2 with probability $1 - \delta_i$ on play n + 1.

T3a: If S_i ($i = 1, 2$) occupies v_3 on play n and both players choose d on play n, S_i will continue to be in v_3 on play n + 1 with probability 1.

T3b: If S_i ($i = 1, 2$) occupies v_3 on play n and the other player chooses c on play n, S_i will be in v_2 with probability γ_i and in v_3 with probability $1 - \gamma_i$ on play n + 1.

Briefly formulated, the parameters can be given the following meaning:

α: the tendency to move into the cooperative state, v_1, after a cooperative reciprocation from the other player.

β: the tendency to move away from the cooperative state, v_1, after a defective response from the other player.

γ: the tendency to move away from the rational state, v_3, after a cooperative response from the other player.

δ: the tendency to move into the rational state, v_3, after a defective reciprocation from the other player.

These transition axioms define for each individual player and for each outcome a matrix of transition probabilities between the value states. These matrices are indicated as $M(cc)$, $M(cd)$, $M(dc)$, and $M(dd)$. For example, $M(cd)$ stands for the transition matrix for S_i, where S_i had chosen c and the other player d. These transition matrices are given in Table 5.2.

It is assumed that each subject begins the sequence of plays in v_2, the transitional state. According to this assumption and the above-mentioned response- and transition-axioms a subject playing in the PDG can be viewed as a probabilistic finite automaton (Paz, 1971). That is, a subject S_i in the PDG is defined as a system (V_i, A, M, (0, 1, 0), $\{v_1\ v_3\}$), where

V_i = the finite set of value states,

A = the finite set of inputs,

M = the set of transition matrices,

(0,1,0) = the initial probability distribution of the value states,

$\{v_1, v_3\}$ = the set of final states.

By the two players of the PDG acting as two automatons a two-actor system is defined, which is an absorbing Markov chain on the ordered pairs of value states of the players. This model is called LAMARC.

The states of LAMARC are elements from the product-set $V_1 \times V_2$. From the assumption that S_1 and S_2 occupy v_2 on play 1 and from the axioms for the single automatoms it follows that LAMARC has 7 (not 9) states. The states are indicated as ordered pairs, where the first symbol represents S_1's value state, and the second symbol S_2's value state.
The states $\langle v_1\ v_3\rangle$ and $\langle v_3\ v_1\rangle$ are excluded, since they can never be entered (as is easy to see). So, 7 latent states remain for LAMARC.

Table 5.2 Set of individual (conditional) transition matrices

Strictly speaking, for each outcome transition probabilities are defined for only two value states. That is, neither in M(*cc*) and M(*cd*) transition probabilities for v_3 are defined, nor are transition probabilities defined for v_1 in M(*dc*) and M(*dd*). However, for practical reasons (for instance the pooling of outcome-specific transition matrices in order to compute the over-all transition matrix) all value states are entered in each transition matrix. Since Prob $(d|v_1) = 0$ and Prob $(c|v_3) = 0$ the added rows and columns act as dummies.

	v_1	v_2	v_3
v_1	1.	0.	0.
v_2	α_i	$1-\alpha_i$	0.
v_3	0.	0.	1.

1) M(*cc*)

	v_1	v_2	v_3
v_1	$1-\beta_i$	β_i	0
v_2	0.	1.	0.
v_3	0.	0.	1.

2) M(*cd*)

	v_1	v_2	v_3
v_1	1.	0.	0.
v_2	0.	1.	0.
v_3	0.	γ_i	$1-\gamma_i$

3) M(*dc*)

	v_1	v_2	v_3
v_1	1.	0.	0.
v_2	0.	$1-\delta_i$	δ_i
v_3	0.	0.	1.

4) M(*dd*)

Table 5.3 shows the conditional outcome probabilities for LAMARC

Table 5.3. Conditional outcome probabilities

	(cc)	(cd)	(dc)	(dd)
$\langle v_1\ v_1 \rangle$	1	0	0	0
$\langle v_1\ v_2 \rangle$	π_2	$(1 - \pi_2)$	0	0
$\langle v_2\ v_1 \rangle$	π_1	0	$(1 - \pi_1)$	0
$\langle v_2\ v_2 \rangle$	$\pi_1\ \pi_2$	$\pi_1(1 - \pi_2)$	$(1 - \pi_1)\pi_2$	$(1 - \pi_1)(1 - \pi_2)$
$\langle v_2\ v_3 \rangle$	0	π_1	0	$(1 - \pi_1)$
$\langle v_3\ v_2 \rangle$	0	0	π_2	$(1 - \pi_2)$
$\langle v_3\ v_3 \rangle$	0	0	0	1

Estimation of the parameters will be performed on massed data of a number of players (from a homogeneous population), Therefore, it is assumed that players have identical parameters.[1)] Figure 5.1a and 5.1b give transition diagrams of LAMARC after having dropped the subscripts

1) Generally, this assumption is unwarranted (see chapter 3). If the data which are pooled are taken from a homogeneous population, this assumption is plausible. As a measure for the homogeneity of the population one can take the variance of the observed outcomes. As we shall see (section 5.3), the data used to evaluate LAMARC showed very small variances for the observed outcomes and therefore can be regarded as being taken from a homogeneous population.

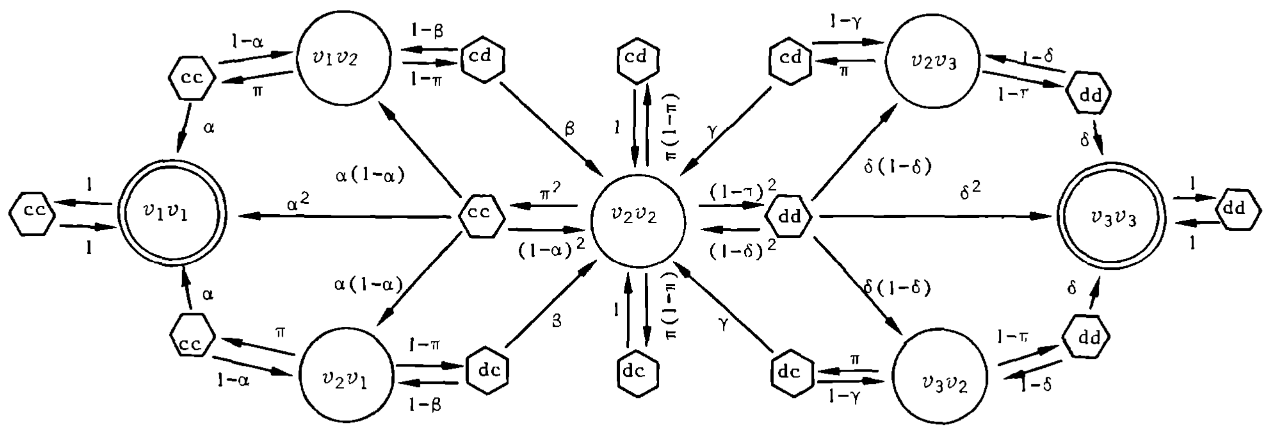


Figure 5.1a: Dual latent states (circles) and response outcomes (hexagons), with response-probabilities and transition probabilities. This figure illustrates tables 5.1, 5.2, and 5.3. Figure 5.1b is the condensation of this figure.

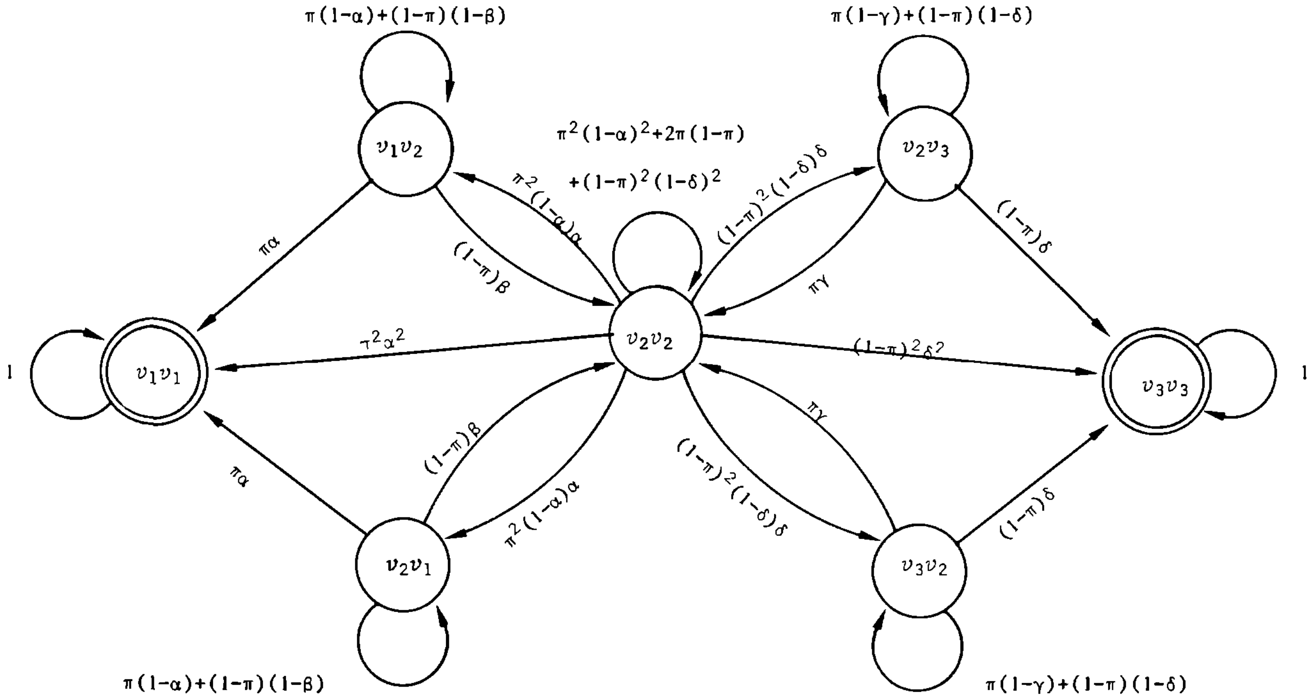


Figure 5.1b: Transition diagram for LAMARC. The final states, which are also absorbing are double-circled.

Table 5.4 Matrix of expected transition probabilities for LAMARC

	$\langle v_1\, v_1 \rangle$	$\langle v_1\, v_2 \rangle$	$\langle v_2\, v_1 \rangle$	$\langle v_2\, v_2 \rangle$	$\langle v_2\, v_3 \rangle$	$\langle v_3\, v_2 \rangle$	$\langle v_3\, v_3 \rangle$
$\langle v_1\, v_1 \rangle$	1	0	0	0	0	0	0
$\langle v_1\, v_2 \rangle$	$\pi\alpha$	$\pi(1-\alpha)+$ $(1-\pi)(1-\beta)$	0	$(1-\pi)\beta$	0	0	0
$\langle v_2\, v_1 \rangle$	$\pi\alpha$	0	$\pi(1-\alpha)+$ $(1-\pi)(1-\beta)$	$(1-\pi)\beta$	0	0	0
$\langle v_2\, v_2 \rangle$	$\pi^2\alpha^2$	$\pi^2(1-\alpha)\alpha$	$\pi^2(1-\alpha)\alpha$	$\pi^2(1-\alpha)^2+2\pi(1-\pi)+$ $(1-\pi)^2(1-\delta)^2$	$(1-\pi)^2(1-\delta)\delta$	$(1-\pi)^2(1-\delta)\delta$	$(1-\pi)^2\delta^2$
$\langle v_2\, v_3 \rangle$	0	0	0	$\pi\gamma$	$\pi(1-\gamma)+$ $(1-\pi)(1-\delta)$	0	$(1-\pi)\delta$
$\langle v_3\, v_2 \rangle$	0	0	0	$\pi\gamma$	0	$\pi(1-\gamma)+$ $(1-\pi)(1-\delta)$	$(1-\pi)\delta$
$\langle v_3\, v_3 \rangle$	0	0	0	0	0	0	1

A subject enters a value state, v_j, as the result of his value state, $v_{j'}$, on the preceding play and the input, $\bar{a} \in A$, i.e. the outcome of the game, on the preceding play: $v_j = f(v_{j'},\bar{a})$. A similar rule holds for the two-player system LAMARC. The expectation of the transition probabilities as indicated in Figure 5.1b are weighted sums (weighted by the conditional outcome probabilities) of the outcome-specific transition probabilities. In Table 5.4 the expectations of the transition probabilities from Figure 5.1b are repeated in matrix form.

If solutions for the parameters of LAMARC are known, expected values can be computed for several non-observable quantities, such as the mean number of plays until absorption is reached in one of the final states, i.e. the mean number of plays until stability is reached, the probability distribution of the absorbing states, etc. The derivation of these statistics is adopted from Snell (1965).

After rearranging the states in Table 5.4 the transition matrix of LAMARC is put in the form.

$$W = \begin{array}{l} \text{absorbing states} \\ \\ \text{transient states} \end{array} \begin{bmatrix} I & 0 \\ & \\ R & Q \end{bmatrix} \qquad (5.1)$$

The entries of the matrix R give the probabilities of moving from the transient states into the absorbing states in one step. The entries of Q give the probabilities of moving in one step from a transient state to the other transient states. The matrix I is the identity matrix of dimension equal to the number of absorbing states, and 0 is the matrix of all 0's with the number of rows equal to the number of absorbing states and the number of columns equal to the number of transient states.

The n-th power of W has the form:

$$W^n = \begin{bmatrix} I & 0 \\ B^{(n)} & Q^n \end{bmatrix} \tag{5.2}$$

The ij-th entry of $B^{(n)}$ gives the probability that starting in the transient state i the process is absorbed in state j during the first n plays. These probabilities are nonincreasing and hence tend to a limiting value which is indicated as b_{ij} $(b_{ij} \neq 0)$.

The ij-th entry of Q^n gives the probability that starting in the transient state i the chain is in the transient state j after n plays. Since the process is eventually absorbed, these probabilities must tend to 0. Thus matrix W^n approaches a limiting matrix W^* of the form

$$W^* = \begin{bmatrix} I & 0 \\ B & 0 \end{bmatrix} \tag{5.3}$$

The components of W^n approach the corresponding components of W^*. $B^{(n)}$ can be written in the form

$$B^{(n)} = (I + Q + Q^2 + \dots + Q^n)R \tag{5.4}$$

A number of descriptive quantities for absorbing Markov chains is obtained by matrix operations on the so called *fundamental matrix* (Kemeny and Shell, 1960, p. 46). This is the matrix, where

$$N = I + Q + Q^2 + \dots + \dots = \sum_{k=0}^{\infty} Q^k, \tag{5.5}$$

that is, N is the matrix when the game is played an infinite number of times.

It can be proved (see for example Kemeny and Shell, 1960), that for $n \to \infty$

$$B = NR \tag{5.6}$$

The entry b_{ij} of B is the probability that starting in state i the process is eventually absorbed in state j.

N is also the inverse of I - Q, that is

$$N(I - Q) = (I - Q)\,N = I \tag{5.7}$$

The ij-th entry of N is the mean number of times that the process is ever in (the transient) state j (counting the initial state) when it is started in i.

Let t be the number of plays to absorption, that is on the t-th play S_1 and S_2 arrive in $\langle v_1\ v_1\rangle$ or in $\langle v_3\ v_3\rangle$. The distribution and moments of t, being a random variable, depend upon the starting state. The expected value of t^k when the process is started in state i is indicated as $g_i^{(k)} = E_i[\,t^k\,]$. The vector $\underline{g}^{(1)}$ gives the mean number of steps to absorption for each transient state as starting state. Thus $\underline{g}^{(k)}$ gives the k-th moment for these numbers. This implies that

$$\underline{g}^{(1)} = N\underline{1}, \tag{5.8}$$

where $\underline{1}$ is a column vector with all antries 1. The element $g_i^{(1)}$ is the sum of the components of the i-th row of N. The k-th moment of t for starting state i is computed as

$$g_i^{(k)} = E_i[\,t^k\,] = \sum_{j \text{ transient}} w_{ij}\, E_j[\,(t+1)^k\,] + \sum_{j \text{ absorbing}} w_{ij}$$

$$= \sum_{j \text{ transient}} w_{ij}\, E_j[\,\sum_{m=0}^{k} \binom{k}{m} t^m\,] + \sum_{j \text{ absorbing}} w_{ij}$$

$$= \sum_{j \text{ transient}} w_{ij}\, E[\,1 + t^k + \sum_{m=1}^{k-1} \binom{k}{m} t^m\,] + \sum_{j \text{ absorbing}} w_{ij}$$

$$= 1 + \sum_{j \text{ transient}} \omega_{ij} E[t^k] + \sum_{j \text{ transient}} \omega_{ij} \sum_{m=1}^{k-1} \binom{k}{m} E_j[t^m]$$

This can be written in vector form as

$$\underline{g}^{(k)} = \underline{1} + Q\underline{g}^{(k)} + \sum_{m=1}^{k-1} \binom{k}{m} Q\underline{g}^{(m)}$$

or

$$(I - Q)\underline{g}^{(k)} = \underline{1} + \sum_{m=1}^{k-1} \binom{k}{m} Q\underline{g}^{(m)}$$

Since $(I - Q)^{-1} = N$,

$$\underline{g}^{(k)} = N[\underline{1} + \sum_{m=1}^{k-1} \binom{k}{m} Q\underline{g}^{(m)}] \tag{5.9}$$

This recursive equation has the following solution (Snell, 1965, p. 445). Let

$$C(k,m) = \sum_{h=1}^{m} (-1)^{h+m} \binom{m}{h} h^k$$

Then

$$\underline{g}^{(k)} = \sum_{m=1}^{k} (-1)^{k+m} C(k,m) N^m \underline{1} \tag{5.10}$$

It appears from (5.10) that all the moments of t can be found

through arithmetic operations on N.
For example, it is found that

$$\underline{g}^{(2)} = N[2N - I]\,\underline{1} \tag{5.11}$$

$$\underline{g}^{(3)} = N[6N^2 - 6N + I]\,\underline{1} \tag{5.12}$$

$$\underline{g}^{(4)} = N[24N^3 - 36N^2 + 14N - I]\,\underline{1} \tag{5.13}$$

The variance of t for initial state i is computed as $g_i^{(2)} - [g_i^{(1)}]^2$.

The probability that t = n (n = 1, 2, ...) for the various transient states as starting states is given by the components of the vector

$$Q^{n-1}\,R\underline{1}\,. \tag{5.14}$$

It was already mentioned that cumulative probabilities are given by $B^{(n)}$ (equation 5.4).

5.3. *A Test of LAMARC*

The model was evaluated with respect to the data reported in Rapoport and Dale (1966). Since no numerical data were reported [1], but only graphs of observed time courses, the "best-fitting" parameters for LAMARC were estimated by trial and error. Data for two populations, males and females, were reported and the aim of this evaluation was to see how adequate LAMARC was in describing the characteristic features of the data for both populations. Another question was how differences between the male and female populations were reflected in the "solutions" for the parameters of LAMARC and whether these differences were psychologically meaningful

1) Inquiries about the raw data were made, but they were not available any more.

The data consisted of the protocols for 49 male pairs and 63 female pairs of players drawn from games which were all of the Prisoner's Dilemma type but with different numerical payoffs. However, payoff matrices differed only by a multiplicative constant. Data from different games were combined, since otherwise the number of pairs from each population would have been too small to yield sufficiently stable data. Time courses over 300 plays were reported of c-choices, (cc)-outcomes, and $(cd + dc)$-outcomes. Also, the variances of the types of outcomes of the game were given.

From the graphs reported in Rapoport and Dale (1966, p.279) numerical values of the data were estimated by eye. Then parameters for LAMARC were chosen to give a "best fit" in the sense that there was a minimum discrepancy between simulated time courses and the data.

It was observed by Rapoport and Dale that for both populations c and d were chosen with equal frequencies on the first play. Therefore, as a first step in estimating the parameters, for both populations π was set at .5. In Table 5.5 the "solutions" of the parameters are given.

Table 5.5. Estimates of the Parameters for the Rapoport-and-Dale Data

	π	α	β	γ	δ
men	.5	.075	.950	.900	.020
women	.5	.040	.950	.940	.050

Figures 5.2 - 5.7 show comparisons between observed time courses from the Rapoport-and-Dale study and those obtained from simulations of LAMARC with the parameters from Table 5.5.

In general, the agreement between observed and simulated curves is fair. Comparisons for women are better, i.e. have a closer over - all agreement between data and simulations, then for men. The simulated c-curve and $(cd + dc)$-curve for men is too high.

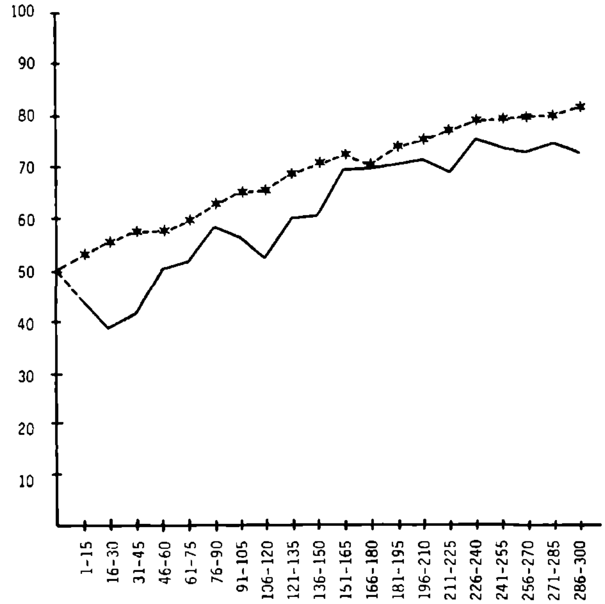


Figure 5.2. Comparison of 49 male pairs and 49 pairs of simulated players for the time course of percentages of *c* choices per block of 15 successive plays. Dashed line: simulation; continuous line: data.

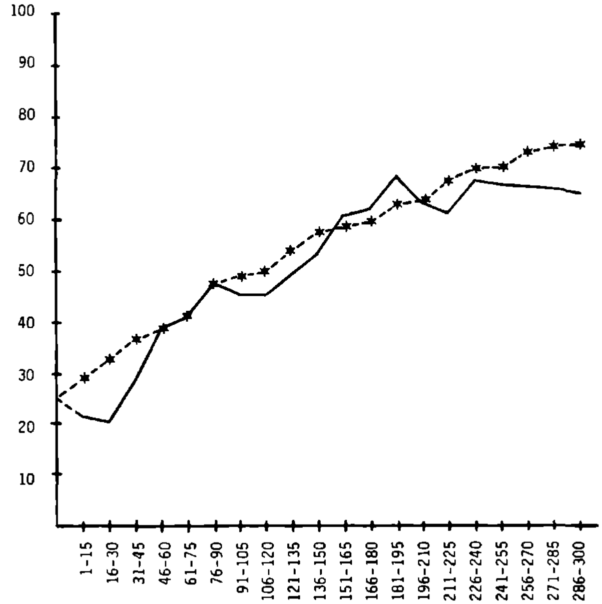


Figure 5.3. Comparisons of 49 male pairs and 49 pairs of simulated players for the time course of percentages of (*cc*) outcomes per block of 15 successive plays. Dashed line: simulation: continuous line: data.

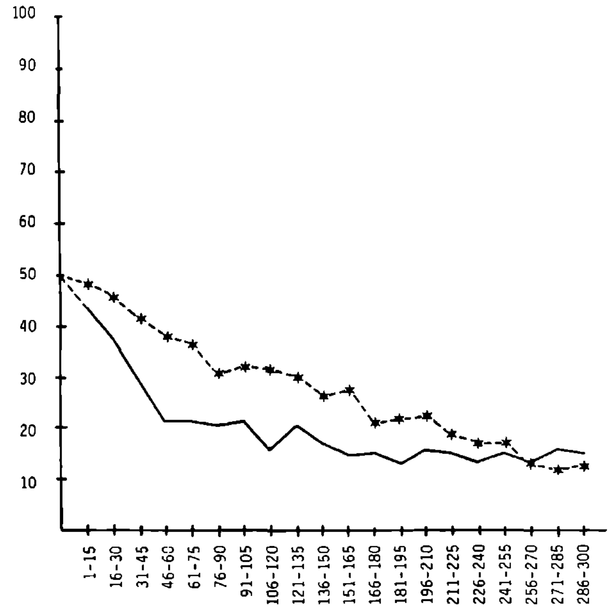


Figure 5.4. Comparison of 49 male pairs and 49 pairs of simulated players for the time course of percentages of (cd + dc) outcomes per block of 15 successive plays. Dashed line: simulation; continuous line: data.

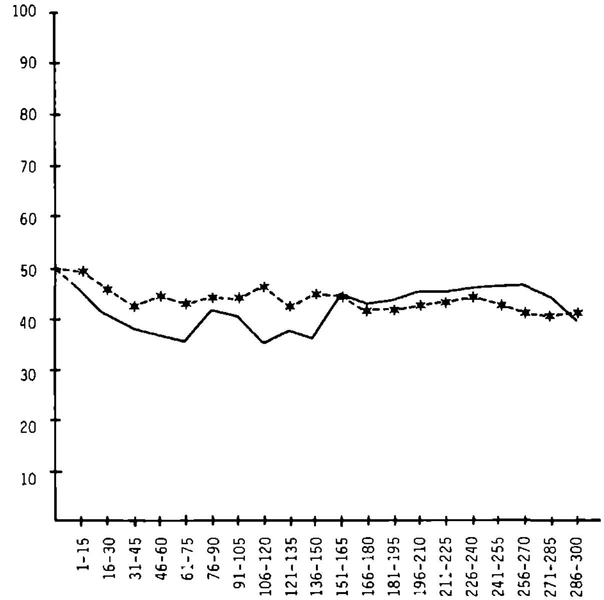


Figure 5.5. Comparisons of 63 female pairs and 63 pairs simulated players for the time course of percentages of c choices per block of 15 successive plays. Dashed line: simulation; continuous line: data.

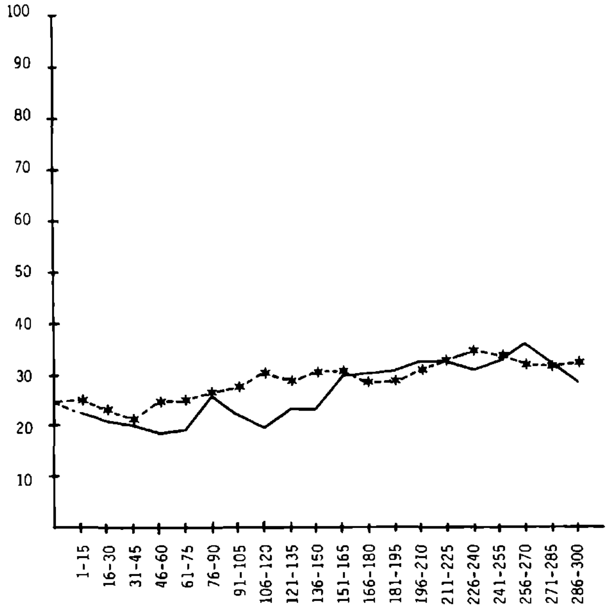


Figure 5.6. Comparisons of 63 female pairs and 63 pairs of simulated players for the time course of percentages of (*cc*) outcomes per block of 15 successive plays. Dashed line: simulation; continuous line: data.

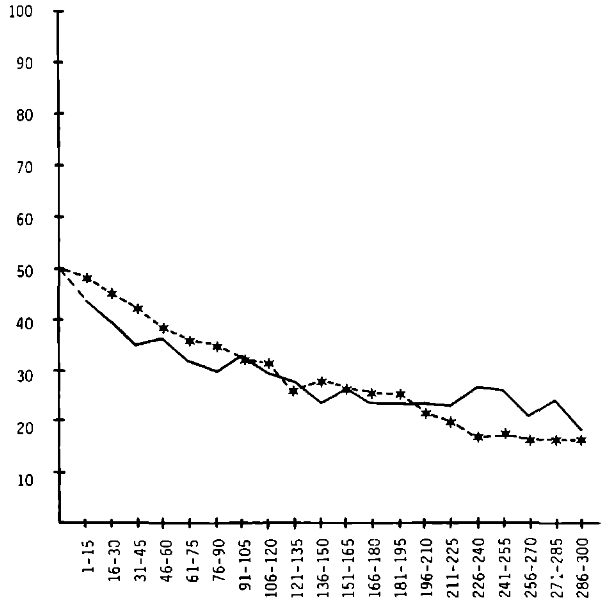


Figure 5.7. Comparison of 63 female pairs and 63 pairs of simulated players for the time course of percentages of (*cd* + *dc*) outcomes per block on 15 successive plays. Dashed line: simulation; continuous line: data.

In Table 5.6 a comparison is made between reported variances of the outcomes for both populations and the simulated players. It is found that there is a good agreement between observed and simulated variances with the comparisons for men having an advantage over those for women.

Table 5.6. Comparison of variances from observed and simulated data for male and female pairs

		(*cc*)	(*cd*)	(*dc*)	(*dd*)
Men	Observed	.110	.007	.008	.056
	Simulated	.080	.008	.008	.029
Women	Observed	.071	.009	.010	.056
	Simulated	.070	.010	.008	.102

From the observed time courses it is seen that women cooperate considerably less than men. How are these differences "explained" by parameters of LAMARC?

Table 5.5 shows that the differences in behavior are reflected in different values for, α, γ, and δ, while β is the same in men and women.

It turns out that these differences range from .00 to .04. An explanation of these differences on statistical grounds cannot be given, since there is neither a measure of the goodness of fit, nor a measure of the significance for the differences found. Therefore, the following explanation does not claim any conclusive force, but instead is very tentative and is given by way of illustration.
The "solutions" suggest that

1. $\alpha_{♂} > \alpha_{♀}$: men decide more easily too cooperate after a double-cooperating outcome, i.e. change more easily to the cooperative state;
2. $\delta_{♀} > \delta_{♂}$: women are more prone to defect after a double-defecting outcome, i.e. change more easily to the rational state;

3. $\gamma_♀ > \gamma_♂$: women are more sensitive to unilateral cooperation by the other player, i.e. leave more easily the rational state.

In general, all subjects tend strongly to move back from either of the "extreme" states (v_1 and v_3) into the "uncertain" state (v_2) after an asymmetric outcome. From this it follows that

- women are more likely to attain stability in (*dd*) and less likely to attain stability in (*cc*) than men,
- it takes more time (that is a greater number of plays) for women to attain stability than it takes for men.

4. Since $\alpha > \delta$ for men and $\delta > \alpha$ for women and π is equal for men and women, men will stabilize in (*cc*) with probability greater than .5 and women will stabilize in (*dd*) with probability greater than .5.

The interpretation of the solutions for the parameters of LAMARC is illustrated with the descriptive quantities derived in the last section. With the solutions for the parameters expected transition matrices can be computed for men and wonen according to the formulas in Table 5.4. These matrices are given in Table 5.7.

Rearrangement of the states in the expected transition matrices yield W according to equation (5.1). From W the submatrix Q is taken and the fundamental matrix N is computed as the inverse matrix of I-Q. The fundamental matrices for male and female pairs are given in Table 5.8. (The symmetries in N follow, of course, from the symmetry of the two players).

Women occupy the uncertein state $\langle v_2, v_3 \rangle$ much more often than men, and are also more often in $\langle v_2, v_3 \rangle$ or $\langle v_3, v_2 \rangle$ than men are, which, of course, follows from the fact that the women change more easily between v_2 and v_3 than the men do, whereas the men move more easily into v_1.

According to the model the process will eventually absorb in one of the two final states, $\langle v_1\ v_1 \rangle$ and $\langle v_3\ v_3 \rangle$. For the observable behavior this means that eventually the frequencies of (*cd*)- and (*dc*)-

Table 5.7 Expected transition matrices of LAMARC for male and female pairs

a) Men

	$\langle v_1\ v_1 \rangle$	$\langle v_1\ v_2 \rangle$	$\langle v_2\ v_1 \rangle$	$\langle v_2\ v_2 \rangle$	$\langle v_2\ v_3 \rangle$	$\langle v_3\ v_2 \rangle$	$\langle v_3\ v_3 \rangle$
$\langle v_1\ v_1 \rangle$	1.	.0	.0	.0	.0	.0	.0
$\langle v_1\ v_2 \rangle$	.0375	.4875	.0	.4750	.0	.0	.0
$\langle v_2\ v_1 \rangle$	.0375	.0	.4875	.4750	.0	.0	.0
$\langle v_2\ v_2 \rangle$	.0014	.0173	.0173	.9540	.0049	.0049	.0001
$\langle v_2\ v_3 \rangle$	.0	.0	.0	.4500	.5400	.0	.0100
$\langle v_3\ v_2 \rangle$	.0	.0	.0	.4500	.0	.5400	.0100
$\langle v_3\ v_3 \rangle$	.0	.0	.0	.0	.0	.0	.1

b) Women

	$\langle v_1\ v_1 \rangle$	$\langle v_1\ v_2 \rangle$	$\langle v_2\ v_1 \rangle$	$\langle v_2\ v_2 \rangle$	$\langle v_2\ v_3 \rangle$	$\langle v_3\ v_2 \rangle$	$\langle v_3\ v_3 \rangle$
$\langle v_1\ v_1 \rangle$	1.	.0	.0	.0	.0	.0	.0
$\langle v_1\ v_2 \rangle$	.0200	.5050	.0	.4750	.0	.0	.0
$\langle v_2\ v_1 \rangle$	.0200	.0	.5050	.4750	.0	.0	.0
$\langle v_2\ v_2 \rangle$	.0004	.0096	.0096	.9560	.0119	.0119	.0006
$\langle v_2\ v_3 \rangle$	.0	.0	.0	.4700	.5050	.0	.0250
$\langle v_3\ v_2 \rangle$	.0	.0	.0	.4700	.0	.5050	.0250
$\langle v_3\ v_3 \rangle$	.0	.0	.0	.0	.0	.0	1.

outcomes will extinguish and that for massed data the proportions of (cc) and (dd) will reach asymptotic values which equal the probabilities that the process is absorbed in $\langle v_1\ v_1 \rangle$ and $\langle v_3\ v_3 \rangle$.

Table 5.8 Fundamental matrices of LAMARC for male and female pairs

a) Men

	$\langle v_1\ v_2\rangle$	$\langle v_2\ v_1\rangle$	$\langle v_2\ v_2\rangle$	$\langle v_2\ v_3\rangle$	$\langle v_3\ v_2\rangle$
$\langle v_1\ v_2\rangle$	9.318	7.367	217.694	2.319	2.319
$\langle v_2\ v_1\rangle$	7.367	9.318	217.694	2.319	2.319
$\langle v_2\ v_2\rangle$	7.949	7.949	234.880	2.502	2.502
$\langle v_2\ v_3\rangle$	7.776	7.776	229.774	4.622	2.448
$\langle v_3\ v_2\rangle$	7.776	7.776	229.774	2.448	4.622

b) Women

	$\langle v_1\ v_2\rangle$	$\langle v_2\ v_1\rangle$	$\langle v_2\ v_2\rangle$	$\langle v_2\ v_3\rangle$	$\langle v_3\ v_2\rangle$
$\langle v_1\ v_2\rangle$	8.270	6.250	322.252	7.747	7.747
$\langle v_2\ v_1\rangle$	6.250	8.270	322.252	7.747	7.747
$\langle v_2\ v_2\rangle$	6.513	6.513	335.821	8.073	8.073
$\langle v_2\ v_3\rangle$	6.184	6.184	318.860	9.686	7.666
$\langle v_3\ v_2\rangle$	6.184	6.184	318.860	7.666	9.686

This probability distribution over the absorbing states is found from the product B = NR (see equation 5.5). In B the probability distributions are given for all transient states as starting states. Since it is assumed that the process is started in $\langle v_2\ v_2 \rangle$, only the third row of B is relevant. The probability distributions of the final states of LAMARC for male and female pairs are given in Table 5.9.

From Table 5.9 it appears that men are more likely to absorb in $\langle v_1\ v_1 \rangle$ (and consequently to "lock-in" at cc), and that women are

Table 5.9. Probability distributions of the final states of LAMARC for male and female pairs, when the process is started in $\langle v_2\ v_2 \rangle$

	$\langle v_1\ v_1 \rangle$	$\langle v_3\ v_3 \rangle$
Men	.926	.074
Women	.395	.605

more likely to absorb in $\langle v_3\ v_3 \rangle$ (and consequently to "lock-in" at *dd*).

The expected number of plays to absorption can be found with equation 5.8. Again only the third element of the vector $\underline{g}^{(1)}$ is relevant. It was found, that the expected number of plays to absorption was 256 for male pairs and 365 for female pairs. The standard deviations were aproximately 254 and 363 for male and female pairs respectively.

5.4. *An estimation procedure for LAMARC*

Probably the estimates of the parameters could have been better (in the sense of yielding a better agreement between simulated and observed behavior sequences), if numerical data would have been available. Also, one must remember that estimating was performed by trial and error. However, now it has been shown that the model can rather accurately describe the observed time courses and since it is psychologically meaningful, it makes sense to think of a better estimation procedure.

As a first step the response parameter π is estimated. For this one needs to known when a player in in v_2, the transitional state. This problem is related to *the identifiability problem.* See for example Greeno and Steiner (1964) and Nahinsky (1973).

Greeno and Steiner (1964) suggest to search for sequences of outcomes which have the effect of resetting the system (recurrent events). That is, if a sequence of outcomes on play j, j+1, ..., j+k (j, k = 1, 2, ...) is a recurrent event, anything which happens before play j+k can be ignored when questions are considered about what happens after play j+k.

Since it is assumed that every player begins the sequence of plays in v_2 and according to axiom T2b, it follows that the players occupy $\langle v_2\ v_2 \rangle$ at least until the first (cc) or (dd) outcome has been chosen. For these plays state $\langle v_2\ v_2 \rangle$ is fully identifiable, until after the first (cc) or (dd) response.

Let m be the number of plays that the players are indentified to occupy $\langle v_2\ v_2 \rangle$.

Let m_1 be the number of (cc) outcomes of these m plays, and let m_2 be the total number of (cd) and (dc) outcomes of these m plays, then a maximum likelihood estimate for π is:

$$\hat{\pi} = \frac{2m_1 + m_2}{2m} \tag{5.15}$$

Suppose, on play j (j = 1, 2, ...) the (dd) outcome is chosen. If on play j+k (k = 1, 2, ...) outcome (cc) is chosen for the first time since play j, it follows that on play j+k the pair of players occupies occupies $\langle v_2\ v_2 \rangle$. Similarly, if (dd) is chosen for the first time on play j+k, after (cc) was chosen on play j the players must occupy $\langle v_2\ v_2 \rangle$ on play j+k (cf. Figure 5.1a).

If outcome (cc) is chosen on play n (n = 1, 2, ...) for the first time since the beginning of the sequence of plays or if (cc) is chosen on play j+k (j, k = 1, 2, ...) for the first time since (dd) has been chosen on play j, one knows for certain that the pair of players occupies $\langle v_2\ v_2 \rangle$ on play n or play j+k respectively. Similarly, if (dd) is chosen on play n for the first time since the beginning of the sequence of plays or if (dd) is chosen on play j+k for the first time since (cc) has been chosen on play j, one knows for certain that the pair of players occupies $\langle v_2\ v_2 \rangle$ on play n or play j+k respectively.

Solving for α and δ is suggested by Nahinsky (1973), who solves for the latent distribution in an outcome-contingent model on trial n by assuming a fixed starting state assuming that solutions are known for trial n-1 for a fixed observable outcome (p. 306-310).

Figure 5.1a can be helpful to see the logic of the following derivations.

Let $X_{(j)}$ be the event that (cc) has been chosen on play j and that the pair of players occupies $\langle v_2\ v_2 \rangle$ on play j. Then

$$\text{Prob } [(cc)_{j+1} | X_{(j)}] = \alpha^2 + 2\pi(1 - \alpha)\alpha + \pi^2 (1 - \alpha)^2 = [(1 - \pi)\alpha + \pi]^2$$

$$\begin{aligned}
\text{Prob } [(cd)_{j+1} | X_{(j)}] &= (1 - \pi)(1 - \alpha)\alpha + \pi(1 - \pi)(1 - \alpha)^2 \\
&= (1 - \pi)(1 - \alpha) [(1 - \pi)\alpha + \pi] \\
\text{Prob } [(dc)_{j+1} | X_{(j)}] &= (1 - \pi)(1 - \alpha)\alpha + \pi(1 - \pi)(1 - \alpha)^2 \qquad (5.16) \\
&= (1 - \pi)(1 - \alpha) [(1 - \pi)\alpha + \pi] \\
\text{Prob } [(dd)_{j+1} | X_{(j)}] &= (1 - \pi)^2(1 - \alpha)^2
\end{aligned}$$

Let m_1 be the number of (cc) outcomes on play j+1 summed over all j, (j = 1, 2, ...), m_2 the number of (cd) outcomes on play j+1, m_3 the number of (dc) outcomes on play j+1, and $m - m_1 - m_2 - m_3$ the number of (dd) outcomes on play j+1. Then the likelihood of these data is expressed as

$$L = [(1 - \pi)\alpha + \pi]^{2m_1+m_2+m_3} * (1 - \pi)^{2m-2m_1-m_2-m_3} * (1 - \alpha)^{2m-2m_1-m_2-m_3} \qquad (5.17)$$

The logarithm of this likelihood function is equal to

$$\begin{aligned}
\text{Log } L = &(2m_1 + m_2 + m_3)\log[(1 - \pi)\alpha + \pi] + \\
&(2m - 2m_1 - m_2 - m_3)\log(1 - \pi) + \qquad (5.18) \\
&(2m - 2m_1 - m_2 - m_3)\log(1 - \alpha)
\end{aligned}$$

Taking the partial derivative to α and setting it equal to zero yields

$\frac{\partial \text{ Log L}}{\partial \alpha}$

$$= \frac{(2m_1 + m_2 + m_3)(1 - \pi)}{(1 - \pi)\alpha + \pi} - \frac{2m - 2m_1 - m_2 - m_3}{1 - \alpha}$$

$$= \frac{(2m_1 + m_2 + m_3)(1 - \pi)(1 - \alpha) - (2m - 2m_1 - m_2 - m_3)[(1 - \pi)\alpha + \pi]}{[(1 - \pi)\alpha + \pi](1 - \alpha)} = 0 \tag{5.19}$$

The solution for α becomes

$$\hat{\alpha} = \frac{2m_1 + m_2 + m_3 - 2m\pi}{2m(1 - \pi)} \tag{5.20}$$

Let $Y_{(k)}$ be the event, that (dd) has been chosen on play k and that the pair of players occupies $\langle v_2\ v_2 \rangle$ on play k. Then

$$\begin{aligned}
\text{Prob } [(cc)_{k+1}|Y_{(k)}] &= \pi^2(1 - \delta)^2 \\
\text{Prob } [(cd)_{k+1}|Y_{(k)}] &= \pi(1 - \delta)\delta + \pi(1 - \pi)(1 - \delta)^2 \\
&= \pi(1 - \delta)[\pi\delta + (1 - \pi)] \\
\text{Prob } [(dc)_{k+1}|Y_{(k)}] &= \pi(1 - \delta)\delta + \pi(1 - \pi)(1 - \delta)^2 \\
&= \pi(1 - \delta)[\pi\delta + (1 - \pi)] \\
\text{Prob } [(dd)_{k+1}|Y_{(k)}] &= \delta^2 + 2(1 - \pi)(1 - \delta)\delta + (1 - \pi)^2(1 - \delta)^2 \\
&= [\pi\delta + (1 - \pi)]^2
\end{aligned} \tag{5.21}$$

Let now m_1 be the number of (cc) outcomes on play k+1 summed over all k (k = 1, 2, ...), m_2 the number of (cd) outcomes on play k + 1, m_3 the number of (dc) outcomes on play k+1, and $m - m_1 - m_2 - m_3$ the number of (dd) outcomes on play k+1. Then the likelihood of these data is

$$L = [\pi\delta + (1 - \pi)]^{2m-2m_1-m_2-m_3} * \pi^{2m_1+m_2+m_3} * (1 - \delta)^{2m_1+m_2+m_3} \tag{5.22}$$

The logarithm of this likelohood fuction is equal to

$$\begin{aligned} \text{Log } L = {} & (2m - 2m_1 - m_2 - m_3)\log[\pi\delta + (1 - \pi)] + \\ & (2m_1 + m_2 + m_3) \log (\pi) + \\ & (2m \;\; + m \;\; + m\;)\log(1 - \delta) \end{aligned} \tag{5.23}$$

Taking the partial derivative to δ and setting it equal to zero yields

$$\frac{\partial \text{ Log } L}{\partial\delta} = \frac{(2m - 2m_1 - m_2 - m_3)\pi(1 - \delta)-(2m_1 + m_2 + m_3)[\pi\delta + (1 - \pi)]}{[\pi\delta + (1 - \pi)\;(1 - \delta)]} = 0 \tag{5.24}$$

The solution for δ becomes

$$\hat{\delta} = \frac{2m\pi - 2m_1 - m_2 - m_3}{2m\pi} \tag{5.25}$$

According to axioms T1a and T1b transition from v_2 to v_2 is only possible after unilateral defection by the other player. This transition probability is indicated as β.

Unlike with v_2 one never knows for certain when a player is in v_1. Transition from v_2 to v_1 is possible after a (cc) outcome.

When a sequence of (cc) outcomes is observed, it can never be determined wether v_1 has been entered by any player. However, sometimes one can tell that v_1 has not been entered yet!

Let $X_{(j)}$ be an outcome as defined before and let $(cd)^i_{j+1}$ be the asymmetrical outcome on play j+1, where S_i has chosen c.

Let c^i_{j+2} be S_i's choice of c on play j+2. We estimate β as follows: Let f_1 be the number of $[c^i,(cd)^i]$ sequences on plays j+2 and j+1, and let f_2 be the number of $[d^i,(cd)^i]$ sequences om plays j+2 and j+1. Be $F = f_1 + f_2$ the number of $(cd)^i$ outcomes on play j+1. Then (cf. Figure 5.1a):

$$\begin{aligned}\mathrm{Prob}[c^i_{j+2},\ (cd)^i_{j+1}|X_{(j)}] &\\ &= \alpha(1-\alpha)(1-\pi)[\beta\pi + (1-\beta)] + (1-\alpha)^2\pi^2(1-\pi). \\ &= (1-\alpha)(1-\pi)[\alpha - \alpha(1-\pi)\beta + (1-\alpha)\pi^2]\end{aligned} \tag{5.26}$$

$$\begin{aligned}\mathrm{Prob}[d^i_{j+2},\ (cd)^i_{j+1}|X_{(j)}] &= \alpha(1-\alpha)(1-\pi)^2\beta + (1-\alpha)^2\pi(1-\pi)^2 \\ &= (1-\alpha)(1-\pi)^2[\alpha\beta + (1-\alpha)\pi]\end{aligned} \tag{5.27}$$

The likelihood function is

$$\begin{aligned}L = (1-\alpha)^F * (1-\pi)^{F+f_2} * [\alpha - \alpha(1-\pi)\beta + \\ (1-\alpha)\pi^2]^{f_1} * [\alpha\beta + (1-\alpha)\pi]^{f_2}\end{aligned} \tag{5.28}$$

$$\frac{\partial \log L}{\partial \beta} = \frac{-f_1\alpha(1-\pi)}{\alpha - \alpha(1-\pi)\beta + (1-\alpha)\pi^2} + \frac{f_2\alpha}{\alpha\beta + (1-\alpha)\pi} = 0 \tag{5.29}$$

Solving for β yields (assuming $\alpha \neq 0$)

$$f_2[\alpha - \alpha(1-\pi)\beta + (1-\alpha)\pi^2] - f_1(1-\pi)[\alpha\beta + (1-\alpha)\pi] = 0$$

$$\alpha(1-\pi)F\beta = f_2\alpha + (1-\alpha)\pi^2 f_2 - (1-\alpha)\pi(1-\pi)f_1$$

$$\hat{\beta} = \frac{(1-\alpha)\pi(\pi F - f_1) + \alpha f_2}{\alpha(1-\pi)F} \tag{5.30}$$

Next, let $Y_{(k)}$ be as defined before, and let now f_1 be the number of $[c^i, (dc)^i]$ sequences on plays k+2 and k+1 and let f_2 be the number of $[d^i, (dc)^i]$ sequences on plays k+2 and k+1. Be $F = f_1 + f_2$ the number of $(dc)^i$ outcomes on play k+1. Then (cf. Figure 5.1a):

$$\begin{aligned}\mathrm{Prob}[c^i_{k+2}, (dc)^i_{k+1} | Y_{(k)}] &= \delta(1-\delta)\pi^2\gamma + (1-\delta)^2\pi^2(1-\pi) \\ &= (1-\delta)\pi^2[\delta\gamma + (1-\delta)(1-\pi)]\end{aligned} \tag{5.31}$$

$$\begin{aligned}&\mathrm{Prob}[d^i_{k+2}, (dc)^i_{k+1} | Y_{(k)}] \\ &\quad = \delta(1-\delta)\pi[\gamma(1-\pi) + (1-\gamma)] + (1-\delta)^2\pi(1-\pi)^2 \\ &\quad = (1-\delta)\pi[\delta - \pi\delta\gamma + (1-\delta)(1-\pi)^2]\end{aligned} \tag{5.32}$$

The likelihood function is

$$\begin{aligned}L = {} & (1-\delta)^F * \pi^{F+f_2} * [\delta\gamma + (1-\delta)(1-\pi)]^{f_1} * [\delta - \pi\delta\gamma + \\ & (1-\delta)(1-\pi)^2]^{f_2}\end{aligned}$$

$$\begin{aligned}\log L = {} & F \log(1-\delta) + (F + f_2)\log\pi + f_1 \log[\delta\gamma + (1-\delta)(1-\pi)] \\ & + f_2 \log[\delta - \pi\delta\gamma + (1-\delta)(1-\pi)^2]\end{aligned}$$

$$\frac{\delta \log L}{\delta\gamma} = \frac{f_1\delta}{\delta\gamma + (1-\delta)(1-\pi)} - \frac{f_2\delta\pi}{\delta - \pi\delta\gamma + (1-\delta)(1-\pi)^2} = 0 \tag{5.33}$$

Solving for γ yields (assuming $\delta \neq 0$):

$$f_1[\delta - \pi\delta\gamma + (1-\delta)(1-\pi)^2] - f_2\pi[\delta\gamma + (1-\delta)(1-\pi)] = 0$$

$$\pi\delta F\gamma = f_1\delta + (1 - \delta)(1 - \pi)^2 f_1 - (1 - \delta)\pi(1 - \pi)f_2$$

$$\hat{\gamma} = \frac{(1 - \delta)(1 - \pi)(f_1 - \tau F) + \delta f_1}{\pi\delta F} \qquad (5.34)$$

This concludes the estimation process

An alternative estimation of β is possible, using (5.26) only, by equating the right member of (5.26) with f_1, i.e. the frequency of $[\, c^i_{j+2}, (cd)^i_{j+1}]$ sequences, given X_j. Then:

$$\frac{f_1}{M} = (1 - \alpha)(1 - \pi)[\alpha\beta(\pi - 1) + \alpha + (1 - \alpha)\pi^2] \qquad (5.35a)$$

and so

$$\hat{\beta} = \frac{(1 - \alpha)(1 - \pi)[\alpha + (1 - \alpha)\pi^2] * M - f_1}{\alpha(1 - \alpha)(1 - \pi)^2 * M}\ , \qquad (5.35b)$$

where M is the number of events X on play j.
Similarly, γ can be estimated by equating f_1, i.e. the frequency of $[\, d^i_{k+2},\ (dc)^i_{k+1}]$ given Y_k with the right member of (5.32). Then,

$$\frac{f_2}{M} = (1 - \delta)\pi[\delta - \delta\gamma\pi + (1 - \delta)(1 - \pi)^2] \qquad (5.36a)$$

and so

$$\hat{\gamma} = \frac{(1 - \delta)\pi[\delta + (1 - \delta)(1 - \pi^2] * M - f_2}{\delta(1 - \delta)\pi^2 * M} \qquad (5.36b)$$

(In these equations, f_2 is defined in the same way as in the derivation of (5.34) and M stands for the number of events Y on play k).

We want to make a note on the proposed estimation procedure. At the outset of this chapter the estimation procedure was already announced as one having maximum likelihood properties. The reader will have noticed that our estimations of the model's parameters are not based on all the information in the data due to the fact that the states of LAMARC are not completely identifiable. Therefore, in general our estimates will not be sufficient, and sufficiency is one of the properties of maximum likelihood estimators. 1) Since maximum likelihood properties cannot be denied in our estimation procedure, it could be labeled 'incomplete maximum likelihood' or 'quasi maximum likelihood'.

5.5. *Discussion*

The analysis of sequences of behavior in iterated plays of non-zero-sum games such as the Prisoner's Dilemma Game, calls for dynamic decision models which describe the latent processes of prediction, commitment and choice. By applying concepts from Howard's theory of metagames the value-state model, which was originally developed in the context of a 100% cooperative and unresponsive opponent (Meeker, 1971), could be expanded to a general model of dynamic decision making in the PDG.

Although the model presented here seems to make sense when applied to the observed time courses in the PDG, the model should be applied to better (and more) data to see how well the proposed estimation procedure works and whether the model could be refined.

When estimating the parameters of LAMARC for the Rapoport-and-Dale data, a test was done with the introduction of a linear operator for π according to the Bush and Mosteller learning model. The results were negative. However, this could be caused by the lack of accurate data.

England (1973, 1975), who applied successfully similar models to

1) More specifically, if parameters have sufficient estimators, then the maximum likelihood estimate will be a sufficient estimate.

two-party-negotiations, suggests to explore other learning models such as Luce's nonlinear β-model (1959) or Restle's hypothesis testing model (1961) for the response parameter.

A mathematically interesting aspect of LAMARC is that the number of latent states (7) exceeds the number of observable states (4). This challanges Nahinsky's (1973) thesis that these numbers have to be equal in order for a unique underlying structure in the data to be identified.

As is usual when a new model is developed, more questions are raised than answered. For example, can the value-state model be usefully applied to other nonzero-sum games? How can the model be expanded or revised to be applicable to n-person dilemmas? Are the parameters of LAMARC related to the payoffs of the PDG or to personality variables of the players in a systematic way? The tentative results with LAMARC may provide an argument to continue research in this erea.

References

Burns, T. and Meeker, L.D. A Mathematical model of Multi-Dimensional Evaluation, Decision Making,and Social Interaction. In: James L. Cochrane and Milan Zeleny (Eds.), *Multiple Criteria Decision Making*. South Carolina: The University of South Carolina Press, 1973.

Bush, R.R. and Mosteller, F. *Stochastic Models for learning*. New York: John Wiley, 1955.

Cohen, B.P. *Conflict and Conformity*. Cambridge, Mass.: M.I.T. Press, 1963.

England, J.L. Mathematical Models of Two-Party Negotiations. *Behavioral Science*, 1973, *18*, 189-197.

England, J.L. Linear Learning Models for Two-Party Negotiations. *Journal of Conflict Resolution*, 1975, *19*, 682-707.

Estes, W.K. Toward a statistical theory of learning. *Psychological Review*, 1950, *57*, 94-107.

Greeno, J.G. and Steiner, Th.E. Markovian Processes with Indentifiable States: General Considerations and Application to All-or-one Learning. *Psychometrika*, 1964, *29*, 309-333.

Hildebrand, P. *Stable outcomes in Two-Person Games*. University of Waterloo (Canada), Ph. D. Thesis, 1973.

Howard, Nigel. The theory of Metagames. *General Systems*, 1966, Vol. XI, 167-186.

Howard, Nigel. *Paradoxes of Rationality: Theory of Metagames and Political Behavior*. M.I.T. Press, 1971.

Howard, Nigel. Prisoner's Dilemma: The "Solution" by General Metagames. *Behavioral Science*. 1976, *21*, 524-531.

Kemeny, J.G. and Snell, J.L. *Finite Markov Chains*. Princeton, N.J.: Van Nostrand, 1960.

Kuhlman, D.M. and Marshello, A.F.J. Individual Differences in Game Motivation as Moderators of Preprogrammed Strategy Effects in Prisoner's Dilemma. *Journal of Personality and Social Psychology*, 1975, *32*, 922-931.

Leik, R.K. and Meeker, B.F. *Mathematical Sociology*. Englewood Cliffs, N.J.: Prentice Hall, Inc., 1975.

Luce, D. *Individual Choice Behavior*. New York: John Wiley, 1959.

McClintock, C.G. and McNeel, S.P. Reward Level and Game Playing Behavior. *Journal of Conflict Resolution*, 1966, *10*, 98-102.

Meeker, B.F. Value Conflict in Social Exchange: a Markov Model. *Journal of Mathematical Psychology*, 1971, *8*, 389,403.

Messick, D.M. and McClintock, C.G. Motivational Bases of Choice in Experimental Games. *Journal of Experimental Social Psychology*, 1968, *4*, 1-25.

Morehous, L.G. Two Motivations for Defection in Prisoner's Dilemma Games. *General Systems*, 1966, Vol. XI, 225-228.

Nahinsky, J.D. Identifiability of a Class of Theories with States Not Directly Observed Over Trials. *Journal of Mathematical Psychology*, 1973, *10*, 296-325.

Nash, J.F. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, U.S.A., 1950, *36*, 48-49.

Ofshe, L. and Ofshe, R. *Utility and Choice in Social Interaction.* Englewood Cliffs, N.J.: Prentice Hall, Inc., 1970.

Paz, A. *Introduction to Probabilistic Automata.* New York: Academic Press, 1971.

Pruitt, D. Motivational Processes in the Decomposed Prisoner's Dilemma Game. *Journal of Personality and Social Psychology,* 1970, 227-238.

Rapoport, A. and Dale, Ph. Models for Prisoner's Dilemma. *Journal of Mathematical Psychology,* 1966, *3*, 269-286.

Restle, F. *Psychology of Judgement and Choice.* New York: John Wiley, 1961.

Snell, J.L. Stochastic Processes. In: R.D. Luce, R.R. Bush and E. Galanter (Eds.) *Handbook of Mathematical Psychology,* Vol. III. New York: John Wiley, 1965.

Thomas, C.S. *Transitions and offered Sets in Two-Person Games.* Ph. D. -Thesis, University of Waterloo (Canada), 1973.

Thomas, C.S. Design and Conduct of Metagame Theoretic Experiments. In. A. Rapoport (Ed.), *Game Theory as a Theory of Conflict Resolution.* Dordrecht-Holland: D. Reidel Publishing Company, 1974.

Van der Sanden, A.L. Latent Markov chain analysis of a value conflict in Prisoner's Dilemma Games. *British Journal of Mathematical and Statistical Psychology,* 1978, *31*, 126-143.

Van der Sanden, A.L. *Stability in the Prisoner's Dilemma an experimental test with general matagames.* Report 79 MA 01, Department of Psychology, University of Nijmegen (The Netherlands), 1979.

Wyer, R.S. jr. Effects of Outcome Matrix and Partner's Behavior in Two-Person Games. *Journal of Experimental Social Psychology,* 1971, *7*, 190-210.

EPILOGUE

Game theory is part of the theory of decision making. The task of the mathematician is to prescribe 'optimal' choices or strategies according to the principle of rationality, the task of the psychologist is to study the process of decision making which precedes actual behavior. In this dissertation an attempt is made to understand gaming behavior in terms of the cognitive operations performed on prediction and expectation about own and other's behavior in the domain of the Prisoner's Dilemma Game.

The relevance of prediction and expectation for a psychological theory of the decision making process in gaming situations has already been stressed by several authors (see for example Pruitt and Kimmel, 1977, Schulz and Hesse, 1978), but methodological problems of how to penetrate into the processes underlying overt choices have not been solved beyond the determination of response- and state-conditioned propensities.

The value state theory presented in this dissertation provides a way to test hypotheses about the dynamics of the (latent) decision making process of subjects involved in gaming situations In chapter 2 it was found that the value state model of Meeker is consistent with Coombs' parameterization of the PDG, in which the game is reformulated as an approach-avoidance conflict with no prescribed solution. The results of chapter 2 also presented some support for the approach taken to study the latent behavior. In this context the following anecdote is interesting.

On the European Mathematical Psychology Group Meeting of September 1975 a paper reporting on the results of chapter 2, was given by the author. The peculiar finding that analysis of variance on the data of these experiments did not yield any significant effect (see Table 2.11), raised some scepticism in some of the participants with regard to the validity of these results. Remember that the analysis of variance was performed on choice proportions,which in our view do not permit a proper analysis of the decision making process. In a recent study by Fox and Guyer (1978) on repeated trials of a four-person Prisoner's Dilemma Game, subjects, males and females, made choices under two information conditions: in one condition the choices were stated publicly, in the other condition the number of subjects choosing 'cooperatively' and 'competitively' was known to each player but the choice made by each particular player was not known by the other players. Analysis of variance on choice proportions yielded 6 non-significant F-ratios (for the main effects and interaction effects for Conditions, Sex and Trials) and only 1 moderately significant (even 5 F-ratios were smaller than 1.0) ... . Yet, effects were found through profile analysis on choice-conditioned propensities!

Much is written about the relation between attitude and decision (see for example Upshaw, 1975). In this dissertation an integration between attitude and (meta-)rational choice in the context of the decision making process in gaming behavior is developed. Value states represent attitudinal orientations of people toward their environment. These orientations incorporate beliefs about the environment and valued goals. In gaming situations these beliefs are equivalent to a player's predictions about his opponent's choices plus own commitments, while valued goals determine policies. The theory of (general) metagames is a static theory which yields equilibrium solutions for games, which can be regarded as the possible end results of the dynamic decision making process. These equilibrium solutions are attained through the players' choosing metastrategies consisting of (commitment, policy)-pairs.

The equilibrium solutions of a game, derived from metagame theory, are the sanctionable outcomes of the game. In LAMARC the number of absorbing states equals the number of sanctionable outcomes in the PDG. It would be interesting to investigate whether the theory developed in this dissertation could be applied to other two-person nonzero-sum games. For instance, in the Game of Chicken (see Figure 1.10, page 24) there are three sanctionable outcomes, namely (R_1C_2), (R_2C_1) and also (R_1C_1) (the latter outcome is not an equilibrium of the basic game). A model, similar to LAMARC, for the behavior in repeated plays of the game, would require three absorbing states (according to the sanctionable outcomes of the game) and at least one transient state.

Considering the number of metastrategies in the $(\text{full})^2$ metagame and in view of the correspondence between value states and metastrategies, we see that the theory of metagames suggests the existence of more value states in the PDG than the number which was distinguished in LAMARC. Introducing a refinement of the value state theory in terms of increasing the number of value states in LAMARC would consist in an unraveling of the transitional state. The experimental method utilized in chapter 4 could be adapted to make (at least partially) identifiable these value states (casu quo metastrategies). A similar technique as the one proposed in chapter 3 could be applied to analyze data on repeated trials, when metastrategies are made visible, for the case of an unresponsive, preprogrammed opponent.

Studies on the PDG and similar paradigms have been criticized for their lack of correspondence with real life situations. Although we take the position that (two-person) games should be viewed primarily as decision models in stead of paradigms for social interaction (for a discussion on this topic see Nemeth, 1972), we agree that n-person dilemmas (Hardin, 1968; Hamburger, 1973; Dawes, 1975) are more appealing as analogies of real life than the two-person equivalents. Introducing more than two players, even if each player has only two choice alternatives, could make a possible value state model terribly complicated and/or could create enormous analytical problems. As a first step we could restrict ourselves to those n-person games for which there is a simple correspondence with two-person games. In the same way as Coombs' parameterization of the PDG was developed for the class of separable games, we could start with games which can be conceived of as the sum of a number of two-person games.

In an n-person dilemma where each of n persons has two choices, c and d, let $D(m)$ be the player's payoff for a d choice when m players choose c and let $C(m)$ be the player's payoff for a c choice when m players choose c. Then, Hamburger (1973) has demonstrated that if a game is characterized by a graph in which $C(m)$ and $D(m)$ are linear functions (of m), the game corresponds to n-1 simultaneous PDG's in which each of the n players plays against each of the n-1 others.

Perhaps the research suggested can shed more light on the dynamics of the decision making process preceding overt behavior in dilemma type conflict situations.

Dawes, R.M. Formal Models of Dilemmas in Social Decision Making. In: M.F. Kaplan and S. Schwartz (Eds.), *Human judgement and decision processes*. New York: Academic Press, 1975.

Fox, J. and Guyer M. "Public" Choice and Cooperation in n-Person Prisoner's Dilemma. *Journal of Conflict Resolution*, 1978, *22*, 469 - 481.

Hamburger, H. N-person prisoner's dilemmas. *Journal of Mathematical Sociology*, 1973, *3*, 27 - 48.

Hardin, G. The tragedy of the commons. *Science*, 1968, *162*, 1243 - 1248.

Nemeth, C. A critical analysis of research utilizing the prisoner's dilemma paradigm for the study of bargaining. In: L. Berkowitz (Ed.), *Advances in experimental social psychology*, vol. 6. New York: Academic Press, 1972.

Pruitt, D.G. and Kimmel, M.J. Twenty years of experimental gaming: critique,synthesis, and suggestions for the future. *Annual Review of Psychology*, 1977, *28*, 363 - 392.

Schulz, U. and Hesse, W. Ein Modell für das Spielverhalten in Prisoner's -Dilemma-Sequenzen unter Berücksichtigung der Erwartungen. *Psychologische Beiträge*, 1978, *20*, 551 - 570. (in German)

Upshaw, H.S. Judgement and Decision Processes in the Formation and Change of Social Attitudes. In: M.F. Kaplan and S. Schwartz (Eds.), *Human judgement and decision processes*. New York: Academic Press, 1975.

SUMMARY

This dissertation is concerned with the decision making process of subjects in social conflicts of the Prisoner's Dilemma Game type. The approach taken focuses on the role of value states, that is representations of principles of social or socially just behavior, as determinants of the decision making behavior. A mathematical model is developed, which describes the role of these value states by means of a latent Markov chain.

Already for quite a time competition and cooperation in situations where the interests of interacting people are completely or partly opposed, have been major themes of interest for psychologists and other social scientists. In particular, research on this topic has been inspired by mathematical game theory.

Chapter 1 is devoted to an evaluation of the relevance of game theory for the study of behavior in social conflicts. Unequivocal prescriptions with respect to the optimal behavior in experimental games are derived from the principle of rationality as long as strictly competitive games are concerned. However, the principle of rationality breaks down when nonzero-sum games are involved. It appears that in games as Prisoner's Dilemma, Chicken, the Battle of the Sexes, etcetera, optimal strategy choices in the sense of minimax or sure-thing principle either do not exist or do not yield optimal outcomes. Besides, specific strategy choices in nonzero-sum games sometimes can be given psychological connotations such as individualism, cooperation, competition, altruism, etcetera. Game theory clearly is successful for psychology as soon as it fails as a normative theory.

Several attempts have been made to develop descriptive or explanatory models for human behavior in experimental games, but were all aimed at the observable behavior. A major thesis of this dissertation is that these, so called, traditional approaches have been unsuccessful because they fail to penetrate into the unobservable, latent process of dynamic decision making in the course of interaction. Variables such as subjects' strategies or their motivational orientations which determine their strategy choices, are inferred at best a posteriori from the course of the manifest behavior.

Chapter 2 reports on an experiment on iterated plays of the Prisoner's Dilemma game where a subject's opponent is non-responsive and 100% cooperative. A theory for the decision making process in this conflict situation is discussed, which distinguishes between the manifest behavior and the latent behavior of the decision making subjects involved. Here the concept of 'value states' is introduced. The mathematical model derived from this theory is a latent Markov chain model with the following features:

1. if a person's behavior satisfies a latent Markov chain, his manifest behavior is not a Markov chain;
2. assumptions about the effect of the manifest behavior (both of the subject and of his opponent) on the value state the subject holds, are reflected in the parameters of the latent Markov chain model and in the parameters for the relationships between the latent states and the manifest behavior.

Examination of the relationships between the model's parameters and a recent parametrization of the Prisoner's Dilemma Game enabled us to attach a clear psychological interpretation of the parameters of the model.

Another major finding of the experiment discussed in chapter 2, was that analysis of variance on choice percentages for different treatment conditions, did not reveal any significant effect,while analysis of the same data according to the value state model yielded different parameter values for different treatment conditions.

In chapter 3 the value state model of the preceding chapter is extended to decision making under uncertainty. Still the opponent is unresponsive to the subject's behavior. The model is also reformulated for the case of $k \geq 2$ latent states.

A special section in this chapter is devoted to problems relating the estimability of the model proposed. It appears that the conventional methods of parameter estimation are not suitable here.

The extended model was subjected to an experimental test. Subjects played the Prisoner's Dilemma Game (for 120 trials in succession) against a preprogrammed, probabilistically playing, and non-responsive stooge.

The experimental results were not in contradiction with the results found in chapter 2. Further it turned out, that the conditional response probabilities of the latent Markov chain model are insensible to variations in the response probabilities of the 'opponent'.

Further extension of the value state theory towards an open game playing situation with all real players required a solution of the paradox of rationality (for nonzero-sum games). This is accomplished in Howard's theory of metagames, in which rationality is reformulated and the concept of 'meta-rationality' is introduced. This theory is described in chapter 4. Metagame theory is a descriptive theory on stable outcomes in experimental games.

The theory of metagames shows how certain outcomes of a game, which are not an equilibrium according to normative game theory, can be stable. Stability is achieved by the players choosing (non-observable) metastrategies. Metastrategies are conditional strategies which a player chooses wittingly or unwittingly and which reflect his expectations and predictions with respect to his opponent's behavior and with respect to his own behavior.

In case of the Prisoner's Dilemma Game the double-cooperating outcome can be stable according to the theory of metagames. This means that the theory of metagames can explain an empirical finding which is in contradiction with classical, normative game theory.

In the original formulation of the theory of metagames no justice was done to the symmetrical relations which exist in experimental games. This difficulty is met in Howard's theory of general metagames. The theory of general metagames too predicts that cooperation can be stable in the Prisoner's Dilemma Game.

An experiment was designed to test the assertions of the theory with respect to the Prisoner's Dilemma Game and to determine empirically the metastrategies of the players involved. For that purpose a special form of the Prisoner's Dilemma, labelled the Extended Prisoner's Dilemma Game, was used, with which metastrategies could be made manifest. So far this method has not been applied elsewhere.

The experimental results shed more light on the conditions under which stability in the game is reached and suggested a relationship between the theory of value states and the theory of metagames: value states reflect a subject's expectation of own and other's behavior, which is also expressed by metastrategies.

Through the application of concepts from metagame theory the 'value state' model, originally developed in the context of a 100% cooperative and unresponsive opponent, is expanded in chapter 5 to a general model of dynamic decision making in the Prisoner's Dilemma Game.

The major features of this model, which describes the latent process of prediction, commitment and choice, are:

1. each player's behavior is described as a probabilistic finite automaton with three latent states;
2. both players involved in the Prisoner's Dilemma Game, jointly behave like a system, which is represented by an absorbing Markov chain with seven latent states, of which two are absorbing;
3. parameter estimation is accomplished through a quasi or incomplete maximum likelihood procedure;
4. with the parameters estimated a number of non-observable quantities can be computed, which are very informative of the latent behavior of the players involved.

The model, LAMARC, is evaluated with respect to its descriptive adequacy of the observed time courses of two sets of already existing data. It appeared that differences in the parameter values for two populations, male and female players, could be given a clear, psychologically meaningful interpretation in terms of the dynamics of the decision making process in the Prisoner's Dilemma Game.

Finally, in an epilogue the preceding chapters are synthesized and some lines are drawn, along which further research could proceed.

Dit proefschrift gaat over het besluitvormingsproces van personen in sociale conflictsituaties van het type Prisoner's Dilemma Game. De aanpak richt zich op de rol van waarde-instellingen (value states), dat wil zeggen uitgangspunten van sociaal of sociaal rechtvaardig gedrag, als determinanten van het besluitvormingsgedrag. Een mathematisch model is ontwikkeld, dat de rol van deze waarde-instellingen beschrijft door middel van een latente Markov keten.

Het thema competitie en samenwerking in situaties waarin de belangen van interacterende personen geheel of gedeeltelijk tegengesteld zijn, heeft reeds lange tijd de aandacht van psychologen en andere sociaal-wetenschappelijke onderzoekers. In het bijzonder is daarbij aansluiting gezocht bij de mathematische speltheorie.

Hoofdstuk 1 is gewijd aan een evaluatie van het belang van de speltheorie voor de studie van gedrag in sociale conflictsituaties. Zolang het gaat om spelen met strikte competitie kan men ondubbelzinnige voorschriften met betrekking tot het optimaal gedrag in experimentele spelen afleiden vanuit het rationaliteitsbeginsel. Het rationaliteitsbeginsel schiet echter te kort, als het gaat om nietnulsom spelen. Het blijkt, dat in spelen zoals het Prisoner's Dilemma, Chicken, de Battle of the Sexes, etcetera, optimale strategie-keuzen ofwel niet bestaan, ofwel niet leiden tot optimale uitkomsten. Daarnaast kunnen aan bepaalde strategie-keuzen in nietnul-som spelen soms psychologische connotaties worden gegeven zoals individualisme, samenwerking, competitie, altruisme, etcetera. Speltheorie kan blijkbaar met succes worden gebruikt in de psychologie zodra hij te kort schiet als een normatieve theorie.

Er zijn verschillende pogingen ondernomen om descriptieve of verklarende modellen voor menselijk gedrag in experimentele spelen te ontwikkelen. Een belangrijke stelling van dit proefschrift is dat deze, zogenaamde, traditionele benaderingen geen succes hebben gehad, omdat zij verzuimen door te dringen in het niet-waarneembare, latente dynamische besluitvormingsproces tijdens het verloop van de interactie. Variabelen zoals de strategieën die subjecten volgen, of hun motivationele instellingen die hun strategie-keuzen bepalen, zijn op zijn hoogst a posteriori uit het verloop van het manifeste gedrag af te leiden.

Hoofdstuk 2 geeft een verslag van een experiment met herhaalde spelen van het Prisoner's Dilemma Game, waarin de tegenspeler van de proefpersoon niet-beïnvloedbaar en 100% cooperatief is. Een theorie voor het besluitvormingsproces in deze conflictsituatie wordt besproken, die een onderscheid maakt tussen het manifeste gedrag en het latente gedrag van de proefpersonen die in deze besluitvorming zijn gewikkeld. Hier wordt het begrip 'value state' (waarde-instellingen) geïntroduceerd. Het mathematisch model dat uit deze theorie is afgeleid, is een latente Markov keten model met de volgende kenmerken:

1. als het gedrag van iemand beantwoordt aan een latente Markov keten, dan is zijn manifest gedrag geen Markov keten;
2. veronderstellingen omtrent het effect van het manifest gedrag (zowel van het subject als zijn tegenspeler) op de 'waarde-instelling' (value state) van het subject worden weerspiegeld in de parameters van het latente Markov keten model en in de parameters voor de relaties tussen de latente toestanden en het manifest gedrag.

Onderzoek naar de relaties tussen de parameters van het model en een recente parameterizering van het Prisoner's Dilemma Game stelde ons in staat een duidelijke psychologische interpretatie te geven aan de parameters van het model.

Een andere belangrijke bevinding van het experiment dat besproken is in hoofdstuk 2, was dat variantie analyse van keuze-percentages voor verschillende behandelingscondities, geen enkel significant resultaat opleverde, terwijl analyse van dezelfde gegevens volgens het value state model verschillende parameterwaarden voor verschillende behandelingscondities gaf.

In hoofdstuk 3 is het value state model van het voorgaande hoofdstuk uitgebreid voor besluitvorming in onzekerheid. Nog steeds reageert de tegenspeler niet op het gedrag van het subject. Het model is tevens geherformuleerd voor het geval er twee of meer latente toestanden zijn.

Een speciale paragraaf in dit hoofdstuk is gewijd aan schattingsproblemen met betrekking tot het voorgestelde model. Het blijkt dat de gebruikelijke methoden voor parameterschatting hier niet geschikt zijn.

Het uitgebreide model werd onderworpen aan een experimentele test. Proefpersonen speelden het Prisoner's Dilemma Game (120 trials achter elkaar) tegen een voorgeprogrammeerde, volgens het toeval spelende, en niet-beïnvloedbare dummy-speler.

De experimentele resultaten waren niet in strijd met de resultaten die gevonden waren in hoofdstuk 2. Verder bleek, dat de voorwaardelijke response waarschijnlijkheden van het latente Markov keten model ongevoelig zijn voor variaties in de response waarschijnlijkheden van de 'tegenspeler'.

Verdere uitbreiding van de value state theorie naar een open spelsituatie met volledig echte spelers, vereiste een oplossing voor de rationaliteitsparadox (voor nietnul-som spelen). Dit wordt bereikt in Howard's theorie van de metaspelen, in welke theorie rationaliteit is geherformuleerd en het begrip 'meta-rationaliteit' wordt ingevoerd. Deze theorie wordt beschreven in hoofdstuk 4. De theorie van de metaspelen is een descriptieve theorie over stabiele uitkomsten in experimentele spelen.

De theorie van de metaspelen laat zien hoe bepaalde uitkomsten van een spel, die geen evenwichtspunten zijn volgens de normatieve speltheorie, toch stabiel kunnen zijn. Stabilitiet wordt bereikt doordat spelers (niet-waarneembare) metastrategieen kiezen. Metastrategieen zijn voorwaardelijke strategieen die een speler bewust of onbewust kiest en die zijn verwachtingen en voorspellingen weerspiegelen met betrekking tot het gedrag van zijn tegenspeler en met betrekking tot zijn eigen gedrag.

In het geval van het Prisoner's Dilemma Game kan de dubbelcooperatieve uitkomst stabiel zijn volgens de theorie van de metaspelen. Dit betekent dat de theorie van de metaspelen een empirisch resultaat dat in tegenspraak is met de klassieke, normatieve speltheorie, kan verklaren.

In de oorspronkelijke formulering van de theorie van de metaspelen werd geen recht gedaan aan de symmetrische relaties in experimentele spelen. Deze moeilijkheid wordt opgelost in Howard's theorie van de algemene metaspelen. De theory van de algemene metaspelen voospelt oòk, dat cooperatie stabiel kan zijn in het Prisoner's Dilemma Game.

Een experiment werd opgezet om de uitspraken van de theorie met betrekking tot het Prisoner's Dilemma Game te toetsen en om empirisch de metastrategieen vast te stellen van de betrokken spelers. Voor dat doel werd een speciale vorm van het Prisoner's Dilemma, aangeduid als het Uitgebreide Prisoner's Dilemma Game, gebruikt, waarmee metastrategieen zichtbaar konden worden gemaakt. Deze methode is tot nu toe niet elders toegepast.

De experimentele resultaten wierpen meer licht op de condities waaronder stabiliteit in het spel wordt bereikt, en gaven een aanwijzing voor een relatie tussen de value state theory en de theorie van de metaspelen: value states (waarde-instellingen) weerspiegelen iemands verwachting van zijn eigen gedrag en dat van de ander, hetgeen ook wordt uitgedrukt door metastrategieen.

Door toepassing van begrippen uit de theorie van de metaspelen is het value state model, dat oorspronkelijk ontwikkeld werd in samenhang met een 100% cooperatieve en niet-beïnvloedbare tegenspeler, in hoofdstuk 5 uitgebreid naar een algemeen model voor dynamische besluitvorming in het Prisoner's Dilemma Game.

De voornaamste kenmerken van dit model dat het latente proces van predictie, commitment en keuze beschrijft, zijn:

1. het gedrag van elke speler wordt beschreven als een probabilistische eindige automaat met drie latente toestanden;
2. beide spelers in het Prisoner's Dilemma Game gedragen zich samen als een systeem dat weergegeven wordt door een absorberende Markov keten met zeven latente toestanden, waarvan twee absorberend zijn;
3. parameter schatting verloopt via een quasi of incomplete maximum likelihood procedure;
4. met behulp van de geschatte parameters kan een aantal niet-waarneembare grootheden worden berekend, die zeer informatief zijn

met betrekking tot het latente gedrag van de betrokken spelers.

Het model, LAMARC, is geëvalueerd op zijn beschrijvende adequaatheid voor het geobserveerde verloop in de tijd van twee sets reeds bestaande data. Het bleek dat verschillen in de parameterwaarden voor twee populaties, mannelijke en vrouwelijke spelers, een duidelijke, psychologisch zinvolle, interpretatie kon worden gegeven in termen van de dynamiek van het besluitvormingsproces in het Prisoner's Dilemma Game.

Ten slotte worden in een epiloog de voorafgaande hoofdstukken samengevat en worden enkele lijnen uitgezet, waarlangs verder onderzoek zich zou kunnen voltrekken.

CURRICULUM VITAE

Antonius (Toine) Lambertus Maria van der Sanden

geboren 9 juni 1949 te 's-Hertogenbosch

3 juni 1967:	eindexamen gymnasium β aan het Gymnasium "Beekvliet" te Sint-Michielsgestel
september 1967:	aanvang studie in de psychologie aan de Katholieke Universiteit te Nijmegen
30 juni 1970:	kandidaatsexamen psychologie
6 december 1973:	doctoraalexamen psychologie (hoofdrichting mathematische psychologie; met uitbreidingen in ontwikkelingspsychologie en psychologische functieleer; bijvak wiskunde)
1 januari 1974:	wetenschappelijk ambtenaar bij de vakgroep mathematische psychologie aan de Katholieke Universiteit te Nijmegen
1 augustus 1974:	wetenschappelijk assistent aan het Psychology Department van de Melbourne University (Australië)
1 februari 1975:	als wetenschappelijk medewerker bij de Organisatie voor Zuiver Wetenschappelijk Onderzoek (ZWO) verbonden aan de vakgroep mathematische psychologie te Nijmegen
1 mei 1977:	chef projectleiding en methodologie bij de hoofdafdeling persoons- en gezins-enquêtes van het Centraal Bureau voor de Statistiek te Heerlen (vanaf 1 september 1978 tevens plaatsvervangend hoofd van deze hoofdafdeling)
1 november 1977:	docent statistiek voor de opleiding MO-Pedagogiek aan de Katholieke Leergangen Tilburg (tot 1 augustus 1979)
vanaf 1 december 1979:	directeur Stafdiensten en Sociale Zaken van het Instituut voor Doven te Sint-Michielsgestel

STELLINGEN

1. Een niet-sanctioneerbare uitkomst in een ordinaal spel *G* is geen evenwichtspunt van *G*. *(Dit proefschrift)*

2. Indien een stochastisch proces zich gedraagt volgens een latente Markov keten, dan gedraagt de observeerbare output van dit proces zich niet volgens een observeerbare Markov keten.

3. De bewering van Dawes (1975), dat het noodzakelijk is om bij de grafische weergave van een N-persoons sociaal dilemma een metriek te specificeren is onjuist.
   Dawes, R.M. Formal models of dilemmas in social decision making. In: M.F. Kaplan and S. Schwartz (Eds.), Human Judgement and Decision Processes. New York: Academic Press, 1975.

4. Guttman's eendimensionale scalogramanalyse is ongeschikt om de cognitieve ontwikkeling volgens de Piagetiaanse stadia te beschrijven. Daarentegen is een meerdimensionale analyse volgens het conjunctieve, c.q. disjunctieve model van Coombs wel bruikbaar.
   Van der Sanden, A.L. The Conjunctive Model as a Means to Measure Developmental Structures. Research Report 75 MA 11. Nijmegen: Psychologisch Laboratorium der Katholieke Universiteit, 1975.

5. Omdat de iteratieve schattingsprocedure in de multidimensionale schaalanalyse volgens het INDSCAL-model gevoelig is voor locale minima, verdient het aanbeveling om vooraf een globaal idee te hebben over de configuratie van de "object-space", en deze te verwerken in de uitgangsconfiguratie voor de INDSCAL-analyse.
   Carroll, J.D. and Wish, M. Models and methods for three-way multidimensional scaling. In: D.H. Krantz, R.C. Atkinson, R.D. Luce, and P. Suppes (Eds.), Contemporary Developments in Mathematical Psychology (Vol. II). San Francisco: Freeman and Co., 1974.

6. De stelling van Van Uden (1977), dat de ritmische organisatie van gesproken taal van belang is voor het taalverwervingsproces, is in overeenstemming met Raaijmakers' (1979) algemene theorie voor het ophalen van informatie uit het lange-termijn geheugen. De bevindingen dat (geschreven) zinnen beter worden onthouden door dove kinderen die zijn opgevoed volgens de gespreksmethode, dan door dove kinderen die zijn opgevoed in gebarentaal of vingerspelling, kunnen worden verklaard als het in staat zijn van de eerste groep, c.q. het niet in staat zijn van de tweede groep, gebruik te maken van *ritmiek als retrieval cue*
Raaijmakers, J.G. Retrieval from long-term store a general theory and mathematical models. Nijmegen, 1979. (doctoral dissertation, University of Nijmegen)
Van Uden, A. A world of language for deaf children. Part I Basic principles. (Third revised edition). Amsterdam Swets & Zeitlinger, 1977.

7. Gezien het toenemend belang van formele theorieen en het gebruik van mathematische modellen in de hedendaagse psychologie, wordt in de pre-kandidaatsopleiding psychologie in het kader van het vak Statistiek ten onrechte meer aandacht besteed aan de toetsingstheorie dan aan de schattingstheorie.

8. Wanneer de verhoudingen van de marginale totalen voor een classificatievariabele in een contingentietabel niet representatief zijn voor de verhoudingen in de populatie, bestaat het gevaar voor vertekening in een analyse volgens een additief model (bijv. regressie-analyse). In dat geval verdient een analyse volgens een multiplicatief model, zoals het log-lineaire model van Goodman, de voorkeur.
Lindsay, J.K. A comparison of additive and multiplicative models for qualitative data Quality and Quantity, 1975, 9, 43-50.

9. Bij de bestrijding van dopinggebruik in de sport zou een zelfde gedragslijn moeten worden gevolgd als bij de bestrijding van druggebruik: niet de gebruiker dient gestraft te worden, maar degene die de doping verstrekt of voor het gebruik ervan verantwoordelijk is.

10. Gelet op de toenemende werkeloosheid onder academici en gezien het vaak zeer specialistische karakter van academische proefschriften, staat tegenwoordig het schrijven van een dissertatie niet zelden gelijk aan het "zichzelf wegpromoveren uit de arbeidsmarkt".

Nijmegen, 15 november 1979 *A.L.M. van der Sanden*